16TH EDITION

STANDARD HANDBOOK FOR ELECTRICAL ENGINEERS

H. Wayne Beaty • Donald G. Fink

McGraw Hill

SECTION 1
UNITS, SYMBOLS, CONSTANTS, DEFINITIONS, AND CONVERSION FACTORS

H. Wayne Beaty
Editor, Standard Handbook for Electrical Engineers;
Senior Member, Institute of Electrical and Electronics Engineers,
Technical assistance provided by Barry N. Taylor,
National Institute of Standards and Technology

CONTENTS

1.1 THE SI UNITS

The units of the quantities most commonly used in electrical engineering (volts, amperes, watts, ohms, etc.) are those of the metric system. They are embodied in the International System of Units (*Système International d'Unités*, abbreviated *SI*). The SI units are used throughout this handbook, in accordance with the established practice of electrical engineering publications throughout the world. Other units, notably the cgs (centimeter-gram-second) units, may have been used in citations in the earlier literature. The cgs electrical units are listed in Table 1-9 with conversion factors to the SI units.

The SI electrical units are based on the mksa (meter-kilogram-second-ampere) system. They have been adopted by the standardization bodies of the world, including the International Electrotechnical Commission (IEC), the American National Standards Institute (ANSI), and the Standards Board of the Institute of Electrical and Electronics Engineers (IEEE). The United States is the only industrialized nation in the world that does not mandate the use of the SI system. Although the U.S. Congress has the constitutional right to establish measuring units, it has never enforced any system. The metric system (now SI) was legalized by Congress in 1866 and is the only legal measuring system, but other non-SI units are legal as well.

Other English-speaking countries adopted the SI system in the 1960s and 1970s. A few major industries converted, but many people resisted—some for very irrational reasons, denouncing it as "un-American." Progressive businesses and educational institutions urged Congress to mandate SI. As a result, in the 1988 Omnibus Trade and Competitiveness Act, Congress established SI as the *preferred* system for U.S. trade and commerce and urged all federal agencies to adopt it by the end of 1992 (or as quickly as possible without undue hardship). SI remains voluntary for private U.S. business. An excellent book, *Metric in Minutes* (Brownridge, 1994), is a comprehensive resource for learning and teaching the metric system (SI).

1.2 CGPM BASE QUANTITIES

Seven quantities have been adopted by the General Conference on Weights and Measures (CGPM[†]) as *base quantities*, that is, quantities that are not derived from other quantities. The base quantities are length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity. Table 1-1 lists these quantities, the name of the SI unit for each, and the standard letter symbol by which each is expressed in the International System (SI).

TABLE 1-1 SI Base Units

Quantity	Unit	Symbol
Length	meter	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Thermodynamic temperature*	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candela	cd

*Celsius temperature is, in general, expressed in degrees Celsius (symbol °C).

The units of the base quantities have been defined by the CGPM as follows:

meter. The length equal to 1 650 763.73 wavelengths in vacuum of the radiation corresponding to the transition between the levels $^2p_{10}$ and $^5d_5$ of the krypton-86 atom (CGPM).

kilogram. The unit of mass; it is equal to the mass of the international prototype of the kilogram (CGPM).

EDITOR'S NOTE: The prototype is a platinum-iridium cylinder maintained at the International Bureau of Weights and Measures, near Paris. The kilogram is approximately equal to the mass of 1000 cubic centimeters of water at its temperature of maximum density.

second. The duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium = 133 atoms (CGPM).

ampere. The constant current that if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 meter apart in vacuum would produce between these conductors a force equal to 2×10^{-7} newton per meter of length (CGPM).

kelvin. The unit of thermodynamic temperature is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water (CGPM).

EDITOR'S NOTE: The zero of the Celsius scale (the freezing point of water) is defined as 0.01 K below the triple point, that is, 273.15 K. See Table 1-27.

mole. That amount of substance of a system that contains as many elementary entities as there are atoms in 0.012 kilogram of carbon-12 (CGPM).

[†]From the initials of its French name, *Conference Générale des Poids et Mesures.*

NOTE: When the mole is used, the elementary entities must be specified. They may be atoms, molecules, ions, electrons, other particles, or specified groups of such particles.

candela. The luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} Hz and that has a radiant intensity in that direction of 1/683 watt per steradian (CGPM).

EDITOR'S NOTE: Until January 1, 1948, the generally accepted unit of luminous intensity was the *international candle*. The difference between the candela and the international candle is so small that only measurements of high precision are affected. The use of the term *candle* is deprecated.

1.3 SUPPLEMENTARY SI UNITS

Two additional SI units, numerics which are considered as dimensionless derived units (see Sec. 1.4), are the radian and the steradian, for the quantities plane angle and solid angle, respectively. Table 1-2 lists these quantities and their units and symbols. The supplementary units are defined as follows:

radian. The plane angle between two radii of a circle that cut off on the circumference an arc equal in length to the radius (CGPM).

steradian. The solid angle which, having its vertex in the center of a sphere, cuts off an area of the surface of the sphere equal to that of a square with sides equal to the radius of the sphere (CGPM).

TABLE 1-2 SI Supplementary Units

Quantity	Unit	Symbol
Plane angle	radian	rad
Solid angle	steradian	sr

1.4 DERIVED SI UNITS

Most of the quantities and units used in electrical engineering fall in the category of SI derived units, that is, units which can be completely defined in terms of the base and supplementary quantities described above. Table 1-3 lists the principal electrical quantities in the SI system and shows their equivalents in terms of the base and supplementary units. The definitions of these quantities, as they appear in the *IEEE Standard Dictionary of Electrical and Electronics Terms* (ANSI/IEEE Std 100-1988), are

hertz. The unit of frequency 1 cycle per second.

newton. The force that will impart an acceleration of 1 meter per second per second to a mass of 1 kilogram.

pascal. The pressure exerted by a force of 1 newton uniformly distributed on a surface of 1 square meter.

joule. The work done by a force of 1 newton acting through a distance of 1 meter.

watt. The power required to do work at the rate of 1 joule per second.

coulomb. The quantity of electric charge that passes any cross section of a conductor in 1 second when the current is maintained constant at 1 ampere.

volt. The potential difference between two points of a conducting wire carrying a constant current of 1 ampere, when the power dissipated between these points is 1 watt.

farad. The capacitance of a capacitor in which a charge of 1 coulomb produces 1 volt potential difference between its terminals.

ohm. The resistance of a conductor such that a constant current of 1 ampere in it produces a voltage of 1 volt between its ends.

siemens (mho). The conductance of a conductor such that a constant voltage of 1 volt between its ends produces a current of 1 ampere in it.

TABLE 1-3 SI Derived Units in Electrical Engineering

	SI unit			
Quantity	Name	Symbol	Expression in terms of other units	Expression in terms of SI base units
Frequency (of a periodic phenomenon)	hertz	Hz	1/s	s^{-1}
Force	newton	N		$m \cdot kg \cdot s^{-2}$
Pressure, stress	pascal	Pa	N/m^2	$m^{-1} \cdot kg \cdot s^{-2}$
Energy, work, quantity of heat	joule	J	$N \cdot m$	$m^2 \cdot kg \cdot s^{-2}$
Power, radiant flux	watt	W	J/s	$m^2 \cdot kg \cdot s^{-3}$
Quantity of electricity, electric charge	coulomb	C	$A \cdot s$	$s \cdot A$
Potential difference, electric potential, electromotive force	volt	V	W/A	$m^2 \cdot kg \cdot s^{-3} \cdot A^{-1}$
Electric capacitance	farad	F	C/V	$m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$
Electric resistance	ohm	Ω	V/A	$m^2 \cdot kg \cdot s^{-3} \cdot A^{-2}$
Conductance	siemens	S	A/V	$m^{-2} \cdot kg^{-1} \cdot s^3 \cdot A^2$
Magnetic flux	weber	Wb	$V \cdot s$	$m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$
Magnetic flux density	tesla	T	Wb/m^2	$kg \cdot s^{-2} \cdot A^{-1}$
Celsius temperature	degree Celsius	°C	K	
Inductance	henry	H	Wb/A	$m^2 \cdot kg \cdot s^{-2} \cdot A^{-2}$
Luminous flux	lumen	lm		$cd \cdot sr$*
Illuminance	lux	lx	lm/m^2	$m^{-2} \cdot cd \cdot sr$*
Activity (of radionuclides)	becquerel	Bq	l/s	s^{-1}
Absorbed dose	gray	Gy	J/kg	$m^2 \cdot s^{-2}$
Dose equivalent	sievert	Sv	J/kg	$m^2 \cdot s^{-2}$

*In this expression, the steradian (sr) is treated as a base unit. See Table 1-2.

weber. The magnetic flux whose decrease to zero when linked with a single turn induces in the turn a voltage whose time integral is 1 volt-second.

tesla. The magnetic induction equal to 1 weber per square meter.

henry. The inductance for which the induced voltage in volts is numerically equal to the rate of change of current in amperes per second.

TABLE 1-4 Examples of SI Derived Units of General Application in Engineering

	SI unit	
Quantity	Name	Symbol
Angular velocity	radian per second	rad/s
Angular acceleration	radian per second squared	rad/s^2
Radiant intensity	watt per steradian	W/sr
Radiance	watt per square meter steradian	$W \cdot m^{-2} \cdot sr^{-1}$
Area	square meter	m^2
Volume	cubic meter	m^3
Velocity	meter per second	m/s
Acceleration	meter per second squared	m/s^2
Wavenumber	1 per meter	m^{-1}
Density, mass	kilogram per cubic meter	kg/m^3
Concentration (of amount of substance)	mole per cubic meter	mol/m^3
Specific volume	cubic meter per kilogram	m^3/kg
Luminance	candela per square meter	cd/m^2

TABLE 1-5 Examples of SI Derived Units Used in Mechanics, Heat, and Electricity

	SI unit		
Quantity	Name	Symbol	Expression in terms of SI base units
Viscosity, dynamic	pascal second	Pa · s	$m^{-1} \cdot kg \cdot s^{-1}$
Moment of force	newton meter	N · m	$m^2 \cdot kg \cdot s^{-2}$
Surface tension	newton per meter	N/m	$kg \cdot s^{-2}$
Heat flux density, irradiance	watt per square meter	W/m^2	$kg \cdot s^{-3}$
Heat capacity	joule per kelvin	J/K	$m^2 \cdot kg \cdot s^{-2} \cdot K^{-1}$
Specific heat capacity, specific entropy	joule per kilogram kelvin	J/(kg · K)	$m^2 \cdot s^{-2} \cdot K^{-1}$
Specific energy	joule per kilogram	J/kg	$m^2 \cdot s^{-2}$
Thermal conductivity	watt per meter kelvin	W/(m · K)	$m \cdot kg \cdot s^{-3} \cdot K^{-1}$
Energy density	joule per cubic meter	J/m^3	$m^{-1} \cdot kg \cdot s^{-2}$
Electric field strength	volt per meter	V/m	$m \cdot kg \cdot s^{-3} \cdot A^{-1}$
Electric charge density	coulomb per cubic meter	C/m^3	$m^{-3} \cdot s \cdot A$
Electric flux density	coulomb per square meter	C/m^2	$m^{-2} \cdot s \cdot A$
Permittivity	farad per meter	F/m	$m^{-3} \cdot kg^{-1} \cdot s^4 \cdot A^2$
Current density	ampere per square meter	A/m^2	
Magnetic field strength	ampere per meter	A/m	
Permeability	henry per meter	H/m	$m \cdot kg \cdot s^{-2} \cdot A^{-2}$
Molar energy	joule per mole	J/mol	$m^2 \cdot kg \cdot s^{-2} \cdot mol^{-1}$
Molar entropy, molar heat capacity	joule per mole kelvin	J/(mol · K)	$m^2 \cdot kg \cdot s^{-2} \cdot K^{-1}mol^{-1}$

lumen. The flux through a unit solid angle (steradian) from a uniform point source of 1 candela; the flux on a unit surface all points of which are at a unit distance from a uniform point source of 1 candela.

lux. The illumination on a surface of 1 square meter on which there is uniformly distributed a flux of 1 lumen; the illumination produced at a surface all points of which are 1 meter away from a uniform point source of 1 candela.

Table 1-4 lists other quantities and the SI derived unit names and symbols useful in engineering applications. Table 1-5 lists additional quantities and the SI derived units and symbols used in mechanics, heat, and electricity.

1.5 SI DECIMAL PREFIXES

All SI units may have affixed to them standard prefixes which multiply the indicated quantity by a power of 10. Table 1-6 lists the standard prefixes and their symbols. A substantial part of the extensive range (10^{36}) covered by these prefixes is in common use in electrical engineering (e.g., gigawatt, gigahertz, nanosecond, and picofarad). The practice of compounding a prefix (e.g., micromicrofarad) is deprecated (the correct term is picofarad).

1.6 USAGE OF SI UNITS, SYMBOLS, AND PREFIXES

Care must be exercised in using the SI symbols and prefixes to follow exactly the capital-letter and lowercase-letter usage prescribed in Tables 1-1 through 1-8, inclusive. Otherwise, serious confusion

TABLE 1-6 SI Prefixes Expressing Decimal Factors

Factor	Prefix	Symbol	Factor	Prefix	Symbol
10^{18}	exa	E	10^{-1}	deci	d
10^{15}	peta	P	10^{-2}	centi	c
10^{12}	tera	T	10^{-3}	milli	m
10^{9}	giga	G	10^{-6}	micro	μ
10^{6}	mega	M	10^{-9}	nano	n
10^{3}	kilo	k	10^{-12}	pico	p
10^{2}	hecto	h	10^{-15}	femto	f
10^{1}	deka	da	10^{-18}	atto	a

may occur. For example, pA is the SI symbol for 10^{-12} of the SI unit for electric current (picoampere), while Pa is the SI symbol for pressure (the pascal).

The spelled-out names of the SI units (e.g., volt, ampere, watt) are not capitalized. The SI letter symbols are capitalized only when the name of the unit stands for or is directly derived from the name of a person. Examples are V for volt, after Italian physicist Alessandro Volta (1745–1827); A for ampere, after French physicist André-Marie Ampère (1775–1836); and W for watt, after Scottish engineer James Watt (1736–1819). The letter symbols serve the function of abbreviations, but they are used without periods.

It will be noted from Tables 1-1, 1-3, and 1-5 that with the exception of the ampere, all the SI electrical quantities and units are derived from the SI base and supplementary units or from other SI derived units. Thus, many of the short names of SI units may be expressed in compound form embracing the SI units from which they are derived. Examples are the volt per ampere for the ohm, the joule per second for the watt, the ampere-second for the coulomb, and the watt-second for the joule. Such compound usage is permissible, but in engineering publications, the short names are customarily used.

Use of the SI prefixes with non-SI units is not recommended; the only exception stated in IEEE Standard 268 is the microinch. Non-SI units, which are related to the metric system but are not decimal multiples of the SI units such as the calorie, torr, and kilogram-force, are specially to be avoided.

A particular problem arises with the universally used units of time (minute, hour, day, year, etc.) that are nondecimal multiples of the second. Table 1-7 lists these and their equivalents in seconds, as well as their standard symbols (see also Table 1-19). The watthour (Wh) is a case in point; it is equal to 3600 joules. The kilowatthour (kWh) is equal to 3 600 000 joules or 3.6 megajoules (MJ). In the mid-1980s, the use of the kilowatthour persisted widely, although eventually it was expected to be replaced by the megajoule, with the conversion factor 3.6 megajoules per kilowatthour. Other aspects in the usage of the SI system are the subject of the following recommendations published by the IEEE:

TABLE 1-7 Time and Angle Units Used in the SI System (Not Decimally Related to the SI Units)

Name	Symbol	Value in SI unit
minute	min	1 min = 60 s
hour	h	1 h = 60 min = 3 600 s
day	d	1 d = 24 h = 86 400 s
degree	°	$1° = (\pi/180)$ rad
minute	′	$1' = (1/60)° = (\pi/10\,800)$ rad
second	″	$1'' = (1/60)' = (\pi/648\,000)$ rad

Frequency. The CGPM has adopted the name *hertz* for the unit of frequency, but cycle per second is widely used. Although cycle per second is technically correct, the name *hertz* is preferred because of the widespread use of cycle alone as a unit of frequency. Use of cycle in place of cycle per second, or kilocycle in place of kilocycle per second, etc., is incorrect.

Magnetic Flux Density. The CGPM has adopted the name *tesla* for the SI unit of magnetic flux density. The name *gamma* shall not be used for the unit nanotesla.

Temperature Scale. In 1948, the CGPM abandoned *centigrade* as the name of the temperature scale. The corresponding scale is now properly named the *Celsius scale*, and further use of *centigrade* for this purpose is deprecated.

Luminous Intensity. The SI unit of luminous intensity has been given the name *candela*, and further use of the old name *candle* is deprecated. Use of the term *candle-power*, either as the name of a quantity or as the name of a unit, is deprecated.

Luminous Flux Density. The common British-American unit of luminous flux density is the lumen per square foot. The name *footcandle*, which has been used for this unit in the United States, is deprecated.

micrometer and micron. The names *micron* for micrometer and *millimicron* for nanometer are deprecated.

gigaelectronvolt (GeV). Because billion means a thousand million in the United States but a million million in most other countries, its use should be avoided in technical writing. The term *billion electronvolts* is deprecated; use gigaelectronvolts instead.

British-American Units. In principle, the number of British-American units in use should be reduced as rapidly as possible. Quantities are not to be expressed in mixed units. For example, mass should be expressed as 12.75 lb, rather than 12 lb or 12 oz. As a start toward implementing this recommendation, the following should be abandoned:

1. British thermal unit (for conversion factors, see Table 1-25).
2. horsepower (see Table 1-26).
3. Rankine temperature scale (see Table 1-27).
4. U.S. dry quart, U.S. liquid quart, and U.K. (Imperial) quart, together with their various multiples and subdivisions. If it is absolutely necessary to express volume in British-American units, the cubic inch or cubic foot should be used (for conversion factors, see Table 1-17).
5. footlambert. If it is absolutely necessary to express luminance in British-American units, the candela per square foot or lumen per steradian square foot should be used (see Table 1-28A).
6. inch of mercury (see Table 1-23C).

1.7 OTHER SI UNITS

Table 1-8 lists units used in the SI system whose values are not derived from the base quantities but from experiment. The definitions of these units, given in the *IEEE Standard Dictionary* (ANSI/IEEE Std 100-1988) are

electronvolt. The kinetic energy acquired by an electron in passing through a potential difference of 1 volt in vacuum.

NOTE: The electronvolt is equal to 1.60218×10^{-19} joule, approximately (see Table 1-25B).

unified atomic mass unit. The fraction 1/12 of the mass of an atom of the nuclide ^{12}C.

NOTE: u is equal to $1.660\,54 \times 10^{-27}$ kg, approximately.

TABLE 1-8 Units Used with the SI System Whose Values Are Obtained Experimentally

Name	Symbol
electronvolt	eV
unified atomic mass unit	u
astronomical unit*	
parsec	pc

*The astronomical unit does not have an international symbol. AU is customarily used in English, UA in French.

astronomical unit. The length of the radius of the unperturbed circular orbit of a body of negligible mass moving around the sun with a sidereal angular velocity of 0.017 202 098 950 radian per day of 86 400 ephemeris seconds.

NOTE: The International Astronomical Union has adopted a value for 1 AU equal to 1.496×10^{11} meters (see Table 1-15C).

parsec. The distance at which 1 astronomical unit subtends an angle of 1 second of arc. 1 pc = 206 264.8 AU = 30 857 × 10^{12} m, approximately (see Table 1-15C).

1.8 CGS SYSTEMS OF UNITS

The units most commonly used in physics and electrical science, from their establishment in 1873 until their virtual abandonment in 1948, are based on the centimeter-gram-second (cgs) electromagnetic and electrostatic systems. They have been used primarily in theoretical work, as contrasted with the SI units (and their "practical unit" predecessors, see Sec. 1.9) used in engineering. Table 1-9 lists the principal cgs electrical quantities and their units, symbols, and equivalent values in SI units. Use of these units in electrical engineering publications has been officially deprecated by the IEEE since 1966.

The cgs units have not been used to any great extent in electrical engineering, since many of the units are of inconvenient size compared with quantities used in practice. For example, the cgs electromagnetic unit of capacitance is the gigafarad.

1.9 PRACTICAL UNITS (ISU)

The shortcomings of the cgs systems were overcome by adopting the volt, ampere, ohm, farad, coulomb, henry, joule, and watt as "practical units," each being an exact decimal multiple of the corresponding electromagnetic cgs unit (see Table 1-9). From 1908 to 1948, the practical electrical units were embodied in the International System Units (ISU, not to be confused with the SI units). During these years, precise formulation of the units in terms of mass, length, and time was impractical because of imprecision in the measurements of the three basic quantities. As an alternative, the units were standardized by comparison with apparatus, called *prototype standards*. By 1948, advances in the measurement of the basic quantities permitted precise standardization by reference to the definitions of the

TABLE 1-9 CGS Units and Equivalents

Quantity	Name	Symbol	Correspondence with SI unit
Electromagnetic system			
Current	abampere	abA	= 10 amperes (exactly)
Voltage	abvolt	abV	= 10^{-8} volt (exactly)
Capacitance	abfarad	abF	= 10^{9} farads (exactly)
Inductance	abhenry	abH	= 10^{-9} henry (exactly)
Resistance	abohm	abΩ	= 10^{-9} ohm (exactly)
Magnetic flux	maxwell	Mx	= 10^{-8} weber (exactly)
Magnetic field strength	oersted	Oe	= 79.577 4 amperes per meter
Magnetic flux density	gauss	G	= 10^{-4} tesla (exactly)
Magnetomotive force	gilbert	Gb	= 0.795 774 ampere
Electrostatic system			
Current	statampere	statA	= 3.335 641 × 10^{-10} ampere
Voltage	statvolt	statV	= 299.792 46 volts
Capacitance	statfarad	statF	= 1.112 650 × 10^{-12} farad
Inductance	stathenry	statH	= 8.987 554 × 10^{11} henrys
Resistance	statohm	statΩ	= 8.987 554 × 10^{11} ohms
Mechanical units			
(equally applicable to the electrostatic and electromagnetic systems)			
Work/energy	erg	erg	= 10^{-7} joule (exactly)
Force	dyne	dyn	= 10^{-5} newton (exactly)

basic units, and the International System Units were officially abandoned in favor of the absolute units. These in turn were supplanted by the SI units which came into force in 1950.

1.10 DEFINITIONS OF ELECTRICAL QUANTITIES

The following definitions are based on the principal meanings listed in the *IEEE Standard Dictionary* (ANSI/IEEE Std 100-1988), which should be consulted for extended meanings, compound terms, and related definitions. The United States Standard Symbols (ANSI/IEEE Std 260, IEEE Std 280) for these quantities are shown in parentheses (see also Tables 1-10 and 1-11). Electrical units used in the United States prior to 1969, with SI equivalents, are listed in Table 1-29.

Admittance (Y). An admittance of a linear constant-parameter system is the ratio of the phasor equivalent of the steady-state sine-wave current or current-like quantity (response) to the phasor equivalent of the corresponding voltage or voltage-like quantity (driving force).

Capacitance (C). Capacitance is that property of a system of conductors and dielectrics which permits the storage of electrically separated charges when potential differences exist between the conductors. Its value is expressed as the ratio of an electric charge to a potential difference.

Coupling Coefficient (k). Coefficient of coupling (used only in the case of resistive, capacitive, and inductive coupling) is the ratio of the mutual impedance of the coupling to the square root of the product of the self-impedances of similar elements in the two circuit loops considered. Unless otherwise specified, coefficient of coupling refers to inductive coupling, in which case $k = M/(L_1L_2)^{1/2}$, where M is the mutual inductance, L_1 the self-inductance of one loop, and L_2 the self-inductance of the other.

Conductance (G)

1. The conductance of an element, device, branch, network, or system is the factor by which the mean-square voltage must be multiplied to give the corresponding power lost by dissipation as heat or as other permanent radiation or as electromagnetic energy from the circuit.
2. Conductance is the real part of admittance.

Conductivity (γ). The conductivity of a material is a factor such that the conduction current density is equal to the electric field strength in the material multiplied by the conductivity.

Current (I). *Current* is a generic term used when there is no danger of ambiguity to refer to any one or more of the currents described below. (For example, in the expression "the current in a simple series circuit," the word *current* refers to the conduction current in the wire of the inductor and to the displacement current between the plates of the capacitor.)

Conduction Current. The conduction current through any surface is the integral of the normal component of the conduction current density over that surface.

Displacement Current. The displacement current through any surface is the integral of the normal component of the displacement current density over that surface.

Current Density (J). *Current density* is a generic term used when there is no danger of ambiguity to refer either to conduction current density or to displacement current density or to both.

Displacement Current Density. The displacement current density at any point in an electric field is (in the International System) the time rate of change of the electric-flux-density vector at that point.

Conduction Current Density. The electric conduction current density at any point at which there is a motion of electric charge is a vector quantity whose direction is that of the flow of positive charge at this point, and whose magnitude is the limit of the time rate of flow of net (positive) charge across a small plane area perpendicular to the motion, divided by this area, as the area taken approaches zero in a macroscopic sense, so as to always include this point. The flow of charge may result from the movement of free electrons or ions but is not in general, except in microscopic studies, taken to include motions of charges resulting from the polarization of the dielectric.

Damping Coefficient (δ). If F is a function of time given by

$$F = A \exp(-\delta t) \sin(2\pi t/T)$$

then δ is the damping coefficient.

Elastance (S). Elastance is the reciprocal of capacitance.

Electric Charge, Quantity of Electricity (Q). Electric charge is a fundamentally assumed concept required by the existence of forces measurable experimentally. It has two forms known as positive and negative. The electric charge on (or in) a body or within a closed surface is the excess of one form of electricity over the other.

Electric Constant, Permittivity of Vacuum (Γ_e). The electric constant pertinent to any system of units is the scalar which in that system relates the electric flux density D in vacuum, to E, the electric field strength ($D = \Gamma_e E$). It also relates the mechanical force between two charges in vacuum to their magnitudes and separation. Thus, in the equation $F = \Gamma_r Q_1 Q_2 / 4\pi \Gamma_e r^2$, the force F between charges Q_1 and Q_2 separated by a distance $r\Gamma_e$ is the electric constant, and Γ_r is a dimensionless factor which is unity in a rationalized system and 4π in an unrationalized system.

NOTE: In the cgs electrostatic system, Γ_e is assigned measure unity and the dimension "numeric." In the cgs electromagnetic system, the measure of Γ_e is that of $1/c^2$, and the dimension is $[L^{-2}T^2]$. In the International System, the measure of Γ_e is $10^7/4\pi c^2$, and the dimension is $[L^{-3}M^{-1}T^4I^2]$. Here, c is the speed of light expressed in the appropriate system of units (see Table 1-12).

*Electric Field Strength (**E**).* The electric field strength at a given point in an electric field is the vector limit of the quotient of the force that a small stationary charge at that point will experience, by virtue of its charge, as the charge approaches zero.

Electric Flux (Ψ). The electric flux through a surface is the surface integral of the normal component of the electric flux density over the surface.

*Electric Flux Density, Electric Displacement (**D**).* The electric flux density is a quantity related to the charge displaced within a dielectric by application of an electric field. Electric flux density at any point in an isotropic dielectric is a vector which has the same direction as the electric field strength, and a magnitude equal to the product of the electric field strength and the permittivity ϵ. In a nonisotropic medium, ϵ may be represented by a tensor and $\boldsymbol{D}$ is not necessarily parallel to $\boldsymbol{E}$.

*Electric Polarization (**P**).* The electric polarization is the vector quantity defined by the equation $\boldsymbol{P} = (\boldsymbol{D} - \Gamma_e \boldsymbol{E})/\Gamma_r$, where $\boldsymbol{D}$ is the electric flux density, Γ_e is the electric constant, $\boldsymbol{E}$ is the electric field strength, and Γ_r is a coefficient that is set equal to unity in a rationalized system and to 4π in an unrationalized system.

Electric Susceptibility (χ_e). Electric susceptibility is the quantity defined by $\chi_e = (\epsilon_r - 1)/\Gamma_r$, where ϵ_r is the relative permittivity and Γ_r is a coefficient that is set equal to unity in a rationalized system and to 4π in an unrationalized system.

Electrization ($\boldsymbol{E}_i$). The electrization is the electric polarization divided by the electric constant of the system of units used.

Electrostatic Potential (V). The electrostatic potential at any point is the potential difference between that point and an agreed-on reference point, usually the point at infinity.

Electrostatic Potential Difference (V). The electrostatic potential difference between two points is the scalar-product line integral of the electric field strength along any path from one point to the other in an electric field, resulting from a static distribution of electric charge.

Impedance (Z). An impedance of a linear constant-parameter system is the ratio of the phasor equivalent of a steady-state sine-wave voltage or voltage-like quantity (driving force) to the phasor equivalent of a steady-state sine-wave current or current-like quantity (response). In electromagnetic radiation, electric field strength is considered the driving force and magnetic field strength the response. In mechanical systems, mechanical force is always considered as a driving force and velocity as a response. In a general sense, the dimension (and unit) of impedance in a given application may be whatever results from the ratio of the dimensions of the quantity chosen as the driving force to the dimensions of the quantity chosen as the response. However, in the types of systems cited above, any deviation from the usual convention should be noted.

Mutual Impedance. Mutual impedance between two loops (meshes) is the factor by which the phasor equivalent of the steady-state sine-wave current in one loop must be multiplied to give the phasor equivalent of the steady-state sine-wave voltage in the other loop caused by the current in the first loop.

Self-impedance. Self-impedance of a loop (mesh) is the impedance of a passive loop with all other loops of the open-circuited network.

Transfer Impedance. A transfer impedance is the impedance obtained when the response is determined at a point other than that at which the driving force is applied.

NOTE: In the case of an electric circuit, the response may be determined in any branch except that which contains the driving force.

Logarithmic Decrement (Λ). If F is a function of time given by

$$F = A \exp(-\delta t) \sin(2\pi t/T)$$

then the logarithmic decrement $\Lambda = T\delta$.

Magnetic Constant, Permeability of Vacuum (Γ_m). The magnetic constant pertinent to any system of units is the scalar which in that system relates the mechanical force between two currents in vacuum to their magnitudes and geometric configurations. For example, the equation for the force F on a length l of two parallel straight conductors of infinite length and negligible circular cross section, carrying constant currents I_1 and I_2 and separated by a distance r in vacuum, is $F = \Gamma_m \Gamma_r I_{12} l/2\pi r$, where Γ_m is the magnetic constant and Γ_r is a coefficient set equal to unity in a rationalized system and to 4π in an unrationalized system.

NOTE: In the cgs electromagnetic system, Γ_m is assigned the magnitude unity and the dimension "numeric." In the cgs electrostatic system, the magnitude of Γ_m is that of $1/c^2$, and the dimension is $[L^{-2}T^2]$. In the International System, Γ_m is assigned the magnitude $4\pi \times 10^{-7}$ and has the dimension $[LMT^{-2}I^{-2}]$.

*Magnetic Field Strength (**H**).* Magnetic field strength is that vector point function whose curl is the current density and which is proportional to magnetic flux density in regions free of magnetized matter.

Magnetic Flux (Φ). The magnetic flux through a surface is the surface integral of the normal component of the magnetic flux density over the surface.

*Magnetic Flux Density, Magnetic Induction (**B**).* Magnetic flux density is that vector quantity which produces a torque on a plane current loop in accordance with the relation $\boldsymbol{T} = IA\boldsymbol{n} \times \boldsymbol{B}$, where $\boldsymbol{n}$ is the positive normal to the loop and A is its area. The concept of flux density is extended to a point inside a solid body by defining the flux density at such a point as that which would be measured in a thin disk-shaped cavity in the body centered at that point, the axis of the cavity being in the direction of the flux density.

*Magnetic Moment (**m**).* The magnetic moment of a magnetized body is the volume integral of the magnetization. The magnetic moment of a loop carrying current I is $\boldsymbol{m} = (1/2)\int \boldsymbol{r} \times d\boldsymbol{r}$, where $\boldsymbol{r}$ is the radius vector from an arbitrary origin to a point on the loop, and where the path of integration is taken around the entire loop.

NOTE: The magnitude of the moment of a plane current loop is IA, where A is the area of the loop. The reference direction for the current in the loop indicates a clockwise rotation when the observer is looking through the loop in the direction of the positive normal.

*Magnetic Polarization, Intrinsic Magnetic Flux density (**J**, $\boldsymbol{B}_i$).* The magnetic polarization is the vector quantity defined by the equation $\boldsymbol{J} = (\boldsymbol{B} - \Gamma_m \boldsymbol{H})/\Gamma_r$, where $\boldsymbol{B}$ is the magnetic flux density, Γ_m is the magnetic constant, $\boldsymbol{H}$ is the magnetic field strength, and Γ_r is a coefficient that is set equal to unity in a rationalized system and to 4π in an unrationalized system.

Magnetic Susceptibility (χ_m). Magnetic susceptibility is the quantity defined by $\chi_m = (\mu_r - 1)/\Gamma_r$, where μ_r is the relative permeability and Γ_r is a coefficient that is set equal to unity in a rationalized system and to 4π in an unrationalized system.

*Magnetic Vector Potential (**A**).* The magnetic vector potential is a vector point function characterized by the relation that its curl is equal to the magnetic flux density and its divergence vanishes.

Magnetization ($\boldsymbol{M}, \boldsymbol{H}_i$). The magnetization is the magnetic polarization divided by the magnetic constant of the system of units used.

Magnetomotive Force ($\boldsymbol{F}_m$). The magnetomotive force acting in any closed path in a magnetic field is the line integral of the magnetic field strength around the path.

Mutual Inductance (M). The mutual inductance between two loops (meshes) in a circuit is the quotient of the flux linkage produced in one loop divided by the current in another loop, which induces the flux linkage.

Permeability. *Permeability* is a general term used to express various relationships between magnetic flux density and magnetic field strength. These relationships are either (1) *absolute permeability* (μ), which in general is the quotient of a change in magnetic flux density divided by the corresponding change in magnetic field strength, or (2) *relative permeability* (μ_r), which is the ratio of the absolute permeability to the magnetic constant.

Permeance (P_m). Permeance is the reciprocal of reluctance.

Permittivity, Capacitivity (ϵ). The permittivity of a homogeneous, isotropic dielectric, in any system of units, is the product of its relative permittivity and the electric constant appropriate to that system of units.

Relative Permittivity, Relative Capacitivity, Dielectric Constant (ϵ_r). The relative permittivity of any homogeneous isotropic material is the ratio of the capacitance of a given configuration of electrodes with the material as a dielectric to the capacitance of the same electrode configuration with a vacuum as the dielectric constant. Experimentally, vacuum must be replaced by the material at all points where it makes a significant change in the capacitance.

Power (P). Power is the time rate of transferring or transforming energy. *Electric power* is the time rate of flow of electrical energy. The *instantaneous electric power* at a single terminal pair is equal to the product of the instantaneous voltage multiplied by the instantaneous current. If both voltage and current are periodic in time, the time average of the instantaneous power, taken over an integral number of periods, is the *active power,* usually called simply the *power* when there is no danger of confusion.

If the voltage and current are sinusoidal functions of time, the product of the rms value of the voltage and the rms value of the current is called the *apparent power;* the product of the rms value of the voltage and the rms value of the in-phase component of the current is the *active power;* and the product of the rms value of the voltage and the rms value of the quadrature component of the current is called the *reactive power*.

The SI unit of instantaneous power and active power is the watt. The germane unit for apparent power is the voltampere and for reactive power is the var.

Power Factor (F_p). Power factor is the ratio of active power to apparent power.

Q. *Q,* sometimes called *quality factor,* is that measure of the quality of a component, network, system, or medium considered as an energy storage unit in the steady state with sinusoidal driving force which is given by

$$Q = \frac{2\pi \times (\text{maximum energy in storage})}{\text{energy dissipated per cycle of the driving force}}$$

NOTE: For single components such as inductors and capacitors, the Q at any frequency is the ratio of the equivalent series reactance to resistance, or of the equivalent shunt susceptance to conductance. For networks that contain several elements and for distributed parameter systems, the Q is generally evaluated at a frequency of resonance. The *nonloaded* Q of a system is the value of Q obtained when only the incidental dissipation of the system elements is present. The *loaded* Q of a system is the value Q obtained when the system is coupled to a device that dissipates energy. The "period" in the expression for Q is that of the driving force, not that of energy storage, which is usually half of that of the driving force.

Reactance (X). Reactance is the imaginary part of impedance.

Reluctance (R_m). Reluctance is the ratio of the magnetomotive force in a magnetic circuit to the magnetic flux through any cross section of the magnetic circuit.

Reluctivity (ν). Reluctivity is the reciprocal of permeability.

Resistance (R)

1. The resistance of an element, device, branch, network, or system is the factor by which the mean-square conduction current must be multiplied to give the corresponding power lost by dissipation as heat or as other permanent radiation or as electromagnetic energy from the circuit.
2. Resistance is the real part of impedance.

Resistivity (ρ). The resistivity of a material is a factor such that the conduction current density is equal to the electric field strength in the material divided by the resistivity.

Self-inductance (L)

1. Self-inductance is the quotient of the flux linkage of a circuit divided by the current in that same circuit which induces the flux linkage. If v = voltage induced, $v = d(Li)/dt$.
2. Self-inductance is the factor L in the $\frac{1}{2}Li^2$ if the latter gives the energy stored in the magnetic field as a result of the current i.

NOTE: Definitions 1 and 2 are not equivalent except when L is constant. In all other cases, the definition being used must be specified. The two definitions are restricted to relatively slow changes in i, that is, to low frequencies, but by analogy with the definitions, equivalent inductances often may be evolved in high-frequency applications such as resonators and waveguide equivalent circuits. Such "inductances," when used, must be specified. The two definitions are restricted to cases in which the branches are small in physical size when compared with a wavelength, whatever the frequency. Thus, in the case of a uniform 2-wire transmission line it may be necessary even at low frequencies to consider the parameters as "distributed" rather than to have one inductance for the entire line.

Susceptance (B). Susceptance is the imaginary part of admittance.

Transfer Function (H). A transfer function is that function of frequency which is the ratio of a phasor output to a phasor input in a linear system.

Transfer Ratio (H). A transfer ratio is a dimensionless transfer function.

Voltage, Electromotive Force (V). The voltage along a specified path in an electric field is the dot product line integral of the electric field strength along this path. As defined, here voltage is synonymous with potential difference only in an electrostatic field.

1.11 DEFINITIONS OF QUANTITIES OF RADIATION AND LIGHT

The following definitions are based on the principal meanings listed in the *IEEE Standard Dictionary* (ANSI/IEEE Std 100-1988), which should be consulted for extended meanings, compound terms, and related definitions. The symbols shown in parentheses are from Table 1-10.

Candlepower. Candlepower is luminous intensity expressed in candelas (term deprecated by IEEE).

Emissivity, Total Emissivity (ϵ). The total emissivity of an element of surface of a temperature radiator is the ratio of its radiant flux density (radiant exitance) to that of a blackbody at the same temperature.

Spectral Emissivity, ϵ(λ). The spectral emissivity of an element of surface of a temperature radiator at any wavelength is the ratio of its radiant flux density per unit wavelength interval (spectral radiant exitance) at that wavelength to that of a blackbody at the same temperature.

Light. For the purposes of illuminating engineering, light is visually evaluated radiant energy.

NOTE 1: Light is psychophysical, neither purely physical nor purely psychological. Light is not synonymous with radiant energy, however restricted, nor is it merely sensation. In a general nonspecialized sense, light is the aspect of radiant energy of which a human observer is aware through the stimulation of the retina of the eye.

NOTE 2: Radiant energy outside the visible portion of the spectrum must not be discussed using the quantities and units of light; it is nonsense to refer to "ultraviolet light" or to express infrared flux in lumens.

Luminance (Photometric Brightness) (L). Luminance in a direction, at a point on the surface of a source, or of a receiver, or on any other real or virtual surface is the quotient of the luminous flux (Φ) leaving, passing through, or arriving at a surface element surrounding the point, propagated in directions defined by an elementary cone containing the given direction, divided by the product of the solid angle of the cone ($d\omega$) and the area of the orthogonal projection of the surface element on a plane perpendicular to the given direction ($dA \cos \theta$). $L = d^2\Phi/[d\omega(da \cos \theta)] = dI/(dA \cos \theta)$. In the defining equation, θ is the angle between the direction of observation and the normal to the surface.

In common usage, the term *brightness* usually refers to the intensity of sensation which results from viewing surfaces or spaces from which light comes to the eye. This sensation is determined in part by the definitely measurable luminance defined above and in part by conditions of observation such as the state of adaptation of the eye. In much of the literature, the term *brightness,* used alone, refers to both luminance and sensation. The context usually indicates which meaning is intended.

Luminous Efficacy of Radiant Flux. The luminous efficacy of radiant flux is the quotient of the total luminous flux divided by the total radiant flux. It is expressed in lumens per watt.

Spectral Luminous Efficacy of Radiant Flux, $K(\lambda)$. Spectral luminous efficacy of radiant flux is the quotient of the luminous flux at a given wavelength divided by the radiant flux at the wavelength. It is expressed in lumens per watt.

Spectral Luminous Efficiency of Radiant Flux. Spectral luminous efficiency of radiant flux is the ratio of the luminous efficacy for a given wavelength to the value at the wavelength of maximum luminous efficacy. It is a numeric.

NOTE: The term *spectral luminous efficiency* replaces the previously used terms *relative luminosity* and *relative luminosity factor.*

Luminous Flux (Φ). Luminous flux is the time rate of flow of light.

Luminous Flux Density at a Surface. Luminous flux density at a surface is luminous flux per unit area of the surface. In referring to flux incident on a surface, this is called *illumination (E)*. The preferred term for luminous flux *leaving* a surface is *luminous exitance (M)*, which has been called *luminous emittance*.

Luminous Intensity (I). The luminous intensity of a source of light in a given direction is the luminous flux proceeding from the source per unit solid angle in the direction considered ($I = d\Phi/d\omega$).

Quantity of Light (Q). Quantity of light (luminous energy) is the product of the luminous flux by the time it is maintained, that is, it is the time integral of luminous flux.

Radiance (L). Radiance in a direction, at a point on the surface, of a source, or of a receiver, or on any other real or virtual surface is the quotient of the radiant flux (P) leaving, passing through, or arriving at a surface element surrounding the point, and propagated in directions defined by an elementary cone containing the given direction, divided by the product of the solid angle of the cone ($d\omega$) and the area of the orthogonal projection of the surface element on a plane perpendicular to the given direction ($dA \cos \theta$). $L = d^2P/d\omega(dA \cos \theta) = dI/(dA \cos \theta)$. In the defining equation, θ is the angle between the normal to the element of the source and the direction of observation.

Radiant Density (w). Radiant density is radiant energy per unit volume.

Radiant Energy (W). Radiant energy is energy traveling in the form of electromagnetic waves.

Radiant Flux Density at a Surface. Radiant flux density at a surface is radiant flux per unit area of the surface. When referring to radiant flux incident on a surface, this is called *irradiance (E)*. The preferred term for radiant flux *leaving* a surface is *radiant exitance (M)*, which has been called *radiant emittance.*

Radiant Intensity (I). The radiant intensity of a source in a given direction is the radiant flux proceeding from the source per unit solid angle in the direction considered ($I = dP/d\omega$).

Radiant Power, Radiant Flux (P). Radiant flux is the time rate of flow of radiant energy.

1.12 LETTER SYMBOLS

Tables 1-10 and 1-11 list the United States Standard letter symbols for quantities and units (ANSI Std Y10.5, ANSI/IEEE Std 260). A *quantity symbol* is a single letter (e.g., I for electric current) specified as to general form of type and modified by one or more subscripts or superscripts when appropriate. A *unit symbol* is a letter or group of letters (e.g., cm for centimeter), or in a few cases, a special sign, that may be used in the place of the name of the unit.

Symbols for quantities are printed in *italic* type, while symbols for units are printed in roman type. Subscripts and superscripts that are letter symbols for quantities or for indices are printed in roman type as follows:

C_p	heat capacity at constant pressure p
a_{ij}, a_{45}	matrix elements
I_i, I_o	input current, output current

For indicating the vector character of a quantity, ***boldface italic*** type is used (e.g., $\boldsymbol{F}$ for force). Ordinary italic type is used to represent the magnitude of a vector quantity.

The product of two quantities is indicated by writing ab. The quotient may be indicated by writing

$$\frac{a}{b}, \qquad a/b, \qquad \text{or} \qquad ab^{-1}$$

If more than one solidus (/) is required in any algebraic term, parentheses must be inserted to remove any ambiguity. Thus, one may write $(a/b)/c$ or a/bc, but not $a/b/c$.

Unit symbols are written in lowercase letters, except for the first letter when the name of the unit is derived from a proper name, and except for a very few that are not formed from letters. When a compound unit is formed by multiplication of two or more other units, its symbol consists of the symbols for the separate units joined by a raised dot (e.g., N · m for newton = meter). The dot may be omitted in the case of familiar compounds such as watthour (Wh) if no confusion would result. Hyphens should not be used in symbols for compound units. Positive and negative exponents may be used with the symbols for units.

When a symbol representing a unit that has a prefix (see Sec. 1.5) carries an exponent, this indicates that the multiple (or submultiple) unit is raised to the power expressed by the exponent.

Examples:

$$2\ \text{cm}^3 = 2(\text{cm})^3 = 2(10^{-2}\ \text{m})^3 = 2 \cdot 10^{-6}\ \text{m}^3$$

$$1\ \text{ms}^{-1} = 1(\text{ms})^{-1} = 1(10^{-3}\ \text{s})^{-1} = 10^3\ \text{s}^{-1}$$

Phasor quantities, represented by complex numbers or complex time-varying functions, are extensively used in certain branches of electrical engineering. The following notation and typography are standard:

	Notation	Remarks
Complex quantity	Z	$Z = \|Z\| \exp(\mathrm{j}\phi)$ $Z = \mathrm{Re}\, Z + \mathrm{j}\, \mathrm{Im}\, Z$
Real part	Re Z, Z'	
Imaginary part	Im Z, Z''	
Conjugate complex quantity	Z^*	$Z^* = \mathrm{Re}\, Z - \mathrm{j}\, \mathrm{Im}\, Z$
Modulus of Z	$\|Z\|$	
Phase of Z, Argument of Z	arg Z	arg $Z = \phi$

TABLE 1-10 Standard Symbols for Quantities

Quantity	Quantity symbol	Unit based on International System	Remarks
Space and time:			
Angle, plane	$\alpha,\beta,\gamma,\theta,\phi,\psi$	radian	Other Greek letters are permitted where no conflict results.
Angle, solid	$\Omega \cdots \omega$	steradian	
Length	l	meter	
Breadth, width	b	meter	
Height	h	meter	
Thickness	d, δ	meter	
Radius	r	meter	
Diameter	d	meter	
Length of path line segment	s	meter	
Wavelength	λ	meter	
Wave number	$\sigma \cdots \tilde{\nu}$	reciprocal meter	$\sigma = 1/\lambda$ The symbol $\tilde{\nu}$ is used in spectroscopy.
Circular wave number Angular wave number	k	radian per meter	$k = 2\pi/\lambda$
Area	$A \cdots S$	square meter	
Volume	V, υ	cubic meter	
Time	t	second	
Period	T	second	
Time constant	$\tau \cdots T$	second	
Frequency	$f \cdots \nu$	second	
Speed of rotation Rotational frequency	n	revolution per second	
Angular frequency	ω	radian per second	$\omega = 2\pi f$
Angular velocity	ω	radian per second	
Complex (angular) frequency Oscillation constant	$p \cdots s$	reciprocal second	$p = -\delta + j\omega$
Angular acceleration	α	radian per second squared	
Velocity	υ	meter per second	
Speed of propagation of electromagnetic waves	c	meter per second	In vacuum, c_0
Acceleration (linear)	a	meter per second squared	
Acceleration of free fall Gravitational acceleration	g	meter per second squared	
Damping coefficient	δ	neper per second	
Logarithmic decrement	Λ	(numeric)	
Attenuation coefficient	α	neper per meter	
Phase coefficient	β	radian per meter	
Propagation coefficient	γ	reciprocal meter	$\gamma = \alpha + j\beta$
Mechanics:			
Mass	m	kilogram	
(Mass) density	ρ	kilogram per cubic meter	Mass divided by volume
Momentum	p	kilogram meter per second	
Moment of inertia	I, J	kilogram meter squared	

TABLE 1-10 Standard Symbols for Quantities (*Continued*)

Quantity	Quantity symbol	Unit based on International System	Remarks
Force	F	newton	
Weight	W	newton	Varies with acceleration of free fall
Weight density	γ	newton per cubic meter	Weight divided by volume
Moment of force	M	newton meter	
Torque	$T \cdots M$	newton meter	
Pressure	p	newton per square meter	The SI name *pascal* has been adopted for this unit.
Normal stress	σ	newton per square meter	
Shear stress	τ	newton per square meter	
Stress tensor	σ	newton per square meter	
Linear strain	ε	(numeric)	
Shear strain	γ	(numeric)	
Strain tensor	ε	(numeric)	
Volume strain	θ	(numeric)	
Poisson's ratio	μ, ν	(numeric)	Lateral contraction divided by elongation
Young's modulus Modulus of elasticity	E	newton per square meter	$E = \sigma/\varepsilon$
Shear modulus Modulus of rigidity	G	newton per square meter	$G = \tau/\gamma$
Bulk modulus	K	newton per square meter	$K = -p/\theta$
Work	W	joule	
Energy	E, W	joule	U is recommended in thermodynamics for internal energy and for blackbody radiation.
Energy (volume) density	w	joule per cubic meter	
Power	P	watt	
Efficiency	η	(numeric)	
Heat:			
Thermodynamic temperature	$T \cdots \Theta$	kelvin	
Temperature Customary temperature	$t \cdots \theta$	degree Celsius	The word *centigrade* has been abandoned as the name of a temperature scale.
Heat	Q	joule	
Internal energy	U	joule	
Heat flow rate	$\Phi \cdots q$	watt	Heat crossing a surface divided by time
Temperature coefficient	α	reciprocal kelvin	
Thermal diffusivity	α	square meter per second	
Thermal conductivity	$\lambda \cdots k$	watt per meter kelvin	
Thermal conductance	G_θ	watt per kelvin	
Thermal resistivity	ρ_θ	meter kelvin per watt	
Thermal resistance	R_θ	kelvin per watt	
Thermal capacitance Heat capacity	C_θ	joule per kelvin	
Thermal impedance	Z_θ	kelvin per watt	
Specific heat capacity	c	joule per kelvin kilogram	Heat capacity divided by mass
Entropy	S	joule per kelvin	
Specific entropy	s	joule per kelvin kilogram	Entropy divided by mass
Enthalpy	H	joule	
Radiation and light:			
Radiant intensity	$I \cdots I_e$	watt per steradian	
Radiant power Radiant flux	$P, \Phi \cdots \Phi_e$	watt	

(*Continued*)

TABLE 1-10 Standard Symbols for Quantities (*Continued*)

Quantity	Quantity symbol	Unit based on International System	Remarks
Radiant energy	$W, Q \cdots Q_e$	joule	The symbol U is used for the special case of blackbody radiant energy
Radiance	$L \cdots L_e$	watt per steradian square meter	
Radiant exitance	$M \cdots M_e$	watt per square meter	
Irradiance	$E \cdots E_e$	watt per square meter	
Luminous intensity	$I \cdots I_v$	candela	
Luminous flux	$\Phi \cdots \Phi_v$	lumen	
Quantity of light	$Q \cdots Q_v$	lumen second	
Luminance	$L \cdots L_v$	candela per square meter	
Luminous exitance	$M \cdots M_v$	lumen per square meter	
Illuminance Illumination	$E \cdots E_v$	lux	
Luminous efficacy†	$K(\lambda)$	lumen per watt	
Total luminous efficacy	K, K_t	lumen per watt	
Refractive index Index of refraction	n	(numeric)	
Emissivity†	$\epsilon(\lambda)$	(numeric)	
Total emissivity	ϵ, ϵ_t	(numeric)	
Absorptance†	$\alpha(\lambda)$	(numeric)	
Transmittance†	$\tau(\lambda)$	(numeric)	
Reflectance†	$\rho(\lambda)$	(numeric)	
Fields and circuits:			
Electric charge Quantity of electricity	Q	coulomb	
Linear density of charge	λ	coulomb per meter	
Surface density of charge	σ	coulomb per square meter	
Volume density of charge	ρ	coulomb per cubic meter	
Electric field strength	$E \cdots K$	volt per meter	
Electrostatic potential Potential difference	$V \cdots \phi$	volt	
Retarded scalar potential	V_r	volt	
Voltage Electromotive force	$V, E \cdots U$	volt	
Electric flux	Ψ	coulomb	
Electric flux density (Electric) displacement	D	coulomb per square meter	
Capacitivity Permittivity Absolute permittivity	ϵ	farad per meter	Of vacuum, ε_v
Relative capacitivity Relative permittivity Dielectric constant	ϵ_r, κ	(numeric)	
Complex relative capacitivity	ϵ_r^*, κ^*	(numeric)	$\epsilon_r^* = \epsilon'_r - j\epsilon''_r$
Complex relative permittivity			ϵ'_r is positive for lossy materials. The complex absolute permittivity ϵ^* is defined in analogous fashion.
Complex dielectric constant			

TABLE 1-10 Standard Symbols for Quantities (*Continued*)

Quantity	Quantity symbol	Unit based on International System	Remarks	
Electric susceptibility	$\chi_e \cdots \epsilon_i$	(numeric)	$\chi_e = \epsilon_r - 1$	MKSA
Electrization	$E_i \cdots K_i$	volt per meter	$E_i = (D/\Gamma_e) - E$	MKSA
Electric polarization	P	coulomb per square meter	$P = D - \Gamma_e E$	MKSA
Electric dipole moment	p	coulomb meter		
(Electric) current	I	ampere		
Current density	$J \cdots S$	ampere per square meter		
Linear current density	$A \cdots \alpha$	ampere per meter	Current divided by the breadth of the conducting sheet	
Magnetic field strength	H	ampere per meter		
Magnetic (scalar) potential Magnetic potential difference	U, U_m	ampere		
Magnetomotive force	$F, F_m \cdots \mathscr{F}$	ampere		
Magnetic flux	Φ	weber		
Magnetic flux density Magnetic induction	B	tesla		
Magnetic flux linkage	Λ	weber		
(Magnetic) vector potential	A	weber per meter		
Retarded (magnetic) vector potential	A_r	weber per meter		
Permeability Absolute permeability	μ	henry per meter	Of vacuum, μ_v	
Relative permeability	μ_r	(numeric)		
Initial (relative) permeability	μ_o	(numeric)		
Complex relative permeability	μ_r*	(numeric)	$\mu_r* = \mu_r' - j\mu_r''$ μ_r'' is positive for lossy materials. The complex absolute permeability μ^* is defined in analogous fashion.	
Magnetic susceptibility	$\chi_m \cdots \mu_i$	(numeric)	$\chi_m = \mu_r - 1$	MKSA
Reluctivity	ν	meter per henry	$\nu = 1/\mu$	
Magnetization	H_i, M	ampere per meter	$H_i = (B/\Gamma_m) - H$	MKSA
Magnetic polarization Intrinsic magnetic flux density	J, B_i	tesla	$J = B - \Gamma_m H$	MKSA
Magnetic (area) moment	m	ampere meter squared	The vector product $\boldsymbol{m} \times \boldsymbol{B}$ is equal to the torque.	
Capacitance	C	farad		
Elastance	S	reciprocal farad	$S = 1/C$	
(Self-) inductance	L	henry		
Reciprocal inductance	Γ	reciprocal henry		
Mutual inductance	L_{ij}, M_{ij}	henry	If only a single mutual inductance is involved, *M* may be used without subscripts.	
Coupling coefficient	$k \cdots \kappa$	(numeric)	$k = L_{ij}(L_i L_j)^{-1/2}$	
Leakage coefficient	σ	(numeric)	$\sigma = 1 - k^2$	
Number of turns (in a winding)	N, n	(numeric)		
Number of phases	m	(numeric)		
Turns ratio	$n \cdots n_*$	(numeric)		

(*Continued*)

TABLE 1-10 Standard Symbols for Quantities (*Continued*)

Quantity	Quantity symbol	Unit based on International System	Remarks
Transformer ratio	a	(numeric)	Square root of the ratio of secondary to primary self-inductance. Where the coefficient of coupling is high, $a \approx n_*$.
Resistance	R	ohm	
Resistivity	ρ	ohm meter	
Volume resistivity			
Conductance	G	siemens	$G = \text{Re } Y$
Conductivity	γ, σ	siemens per meter	$\gamma = 1/\rho$ The symbol σ is used in field theory, as γ is there used for the propagation coefficient.
Reluctance	$R, R_m \cdots \mathcal{R}$	reciprocal henry	Magnetic potential difference divided by magnetic flux
Permeance	$P, P_m \cdots \mathcal{P}$	henry	$P_m = 1/R_m$
Impedance	Z	ohm	
Reactance	X	ohm	
Capacitive reactance	X_C	ohm	For a pure capacitance, $X_C = -1/\omega C$
Inductive reactance	X_L	ohm	For a pure capacitance, $X_L = \omega L$
Quality factor	Q	(numeric)	See Q in Sec. 1.10.
Admittance	Y	siemens	$Y = 1/Z = G + jB$
Susceptance	B	siemens	$B = \text{Im } Y$
Loss angle	δ	radian	$\delta = (R/\|X\|)$
Active power	P	watt	
Reactive power	$Q \cdots P_q$	var	
Apparent power	$S \cdots P_s$	voltampere	
Power factor	$\cos \phi \cdots F_p$	(numeric)	
Reactive factor	$\sin \phi \cdots F_q$	(numeric)	
Input power	P_i	watt	
Output power	P_o	watt	
Poynting vector	S	watt per square meter	
Characteristic impedance	Z_o	ohm	
Surge impedance			
Intrinsic impedance of a medium	η	ohm	
Voltage standing-wave ratio	S	(numeric)	
Resonance frequency	f_r	hertz	
Critical frequency	f_c	hertz	
Cutoff frequency			
Resonance angular frequency	ω_r	radian per second	
Critical angular frequency	ω_c	radian per second	
Cutoff angular frequency			
Resonance wavelength	λ_r	meter	
Critical wavelength	λ_c	meter	
Cutoff wavelength			
Wavelength in a guide	λ_g	meter	
Hysteresis coefficient	k_h	(numeric)	
Eddy-current coefficient	k_e	(numeric)	
Phase angle	ϕ, θ	radian	
Phase difference			

†(λ) is not part of the basic symbol but indicates that the quantity is a function of wavelength.

TABLE 1-11 Standard Symbols for Units

Unit	Symbol	Notes
ampere	A	SI unit of electric current
ampere (turn)	A	SI unit of magnetomotive force
ampere-hour	Ah	Also A · h
ampere per meter	A/m	SI unit of magnetic field strength
angstrom	Å	1 Å = 10^{-10} m. Deprecated.
atmosphere, standard	atm	1 atm = 101 325 Pa. Deprecated.
atmosphere, technical	at	1 at = 1 kgf/cm^2. Deprecated.
atomic mass unit (unified)	u	The (unified) atomic mass unit is defined as one-twelfth of the mass of an atom of the ^{12}C nuclide. Use of the old atomic mass (amu), defined by reference to oxygen, is deprecated.
atto	a	SI prefix for 10^{-18}
attoampere	aA	
bar	bar	1 bar = 100 kPa. Use of the bar is strongly discouraged, except for limited use in meteorology.
barn	b	1 b = 10^{-28} m^2
barrel	bb1	1 bb1 = 42 gal$_{US}$ = 158.99 L
barrel per day	bb1/d	This is the standard barrel used for petroleum, etc. A different standard barrel is used for fruits, vegetables, and dry commodities.
baud	Bd	In telecommunications, a unit of signaling speed equal to one element per second. The signaling speed in bauds is equal to the reciprocal of the signal element length in seconds.
bel	B	
becquerel	Bq	SI unit of activity of a radionuclide
billion electronvolts	GeV	The name *gigaelectronvolt* is preferred for this unit.
bit	b	In information theory, the bit is a unit of information content equal to the information content of a message, the *a priori* probability of which is one-half. In computer science, the bit is a unit of storage capacity. The capacity, in bits, of a storage device is the logarithm to the base two of the number of possible states of the device.
bit per second	b/s	
British thermal unit	Btu	
calorie (International Table calorie)	cal$_{IT}$	1 cal$_{IT}$ = 4.1868 J. Deprecated.
calorie (thermochemical calorie)	cal	1 cal = 4.1840 J. Deprecated.
candela	cd	SI unit of luminous intensity
candela per square inch	cd/in^2	Use of the SI unit, cd/m^2, is preferred.
candela per square meter	cd/m^2	SI unit of luminance. The name *nit* is sometimes used for this unit.
candle	cd	The unit of luminous intensity has been given the name *candela;* use of the name *candle* for this unit is deprecated.
centi	c	SI prefix for 10^{-2}
centimeter	cm	
centipoise	cP	1 cP = mPa · s. The name centipoise is deprecated.
centistokes	cSt	1 cSt = 1mm^2/s. The name centistokes is deprecated.
circular mil	cmil	1 cmil = $(\pi/4) \cdot 10^{-6}$ in^2
coulomb	C	SI unit of electric charge
cubic centimeter	cm^3	
cubic foot	ft^3	
cubic foot per minute	ft^3/min	
cubic foot per second	ft^3/s	
cubic inch	in^3	
cubic meter	m^3	
cubic meter per second	m^3/s	
cubic yard	yd^3	

(*Continued*)

TABLE 1-11 Standard Symbols for Units (*Continued*)

Unit	Symbol	Notes
curie	Ci	A unit of activity of radionuclide. Use of the SI unit, the becquerel, is preferred, 1 Ci = 3.7×10^{10} Bq.
cycle	c	
cycle per second	Hz, c/s	See hertz. The name *hertz* is internationally accepted for this unit; the symbol Hz is preferred to c/s.
darcy	D	1 D = 1 cP (cm/s) (cm/atm) = 0.986 923 μm^2. A unit of permeability of a porous medium. By traditional definition, a permeability of one darcy will permit a flow of 1 cm^3/s of fluid of 1 cP viscosity through an area of 1 cm^2 under a pressure gradient of 1 atm/cm. For nonprecision work, 1 D may be taken equal to 1 μm^2 and 1 mD equal to 0.001 μm^2. Deprecated.
day	d	
deci	d	SI prefix for 10^{-1}
decibel	dB	
degree (plane angle)	. . . °	
degree (temperature):		
degree Celsius	°C	SI unit of Celsius temperature. The degree Celsius is a special name for the kelvin, for use in expressing Celsius temperatures or temperature intervals.
degree Fahrenheit	°F	Note that the symbols for °C, °F, and °R comprise two elements, written with no space between the ° and the letter that follows. The two elements that make the complete symbol are not to be separated.
degree Kelvin		See kelvin
degree Rankine	°R	
deka	da	SI prefix for 10
dyne	dyn	Deprecated.
electronvolt	eV	
erg	erg	Deprecated.
exa	E	SI prefix for 10^{18}
farad	F	SI unit of capacitance
femto	f	SI prefix for 10^{-15}
femtometer	fm	
foot	ft	
conventional foot of water	ftH_2O	1 ftH_2O = 2989.1 Pa (ISO)
foot per minute	ft/min	
foot per second	ft/s	
foot per second squared	ft/s^2	
foot pound-force	ft · lbf	
footcandle	fc	1 fc = 1 lm/ft^2. The name *lumen per square foot* is also used for this unit. Use of the SI unit of illuminance, the lux (lumen per square meter), is preferred.
footlambert	fL	1 fL = $(1/\pi)$ cd/ft^2. A unit of luminance. One lumen per square foot leaves a surface whose luminance is one footlambert in all directions within a hemisphere. Use of the SI unit, the candela per square meter, is preferred.
gal	Gal	1 Gal = 1 cm/s^2. Deprecated.
gallon	gal	1 gal_{UK} = 4.5461 L 1 gal_{US} = 231 in^3 = 3.7854 L
gauss	G	The gauss is the electromagnetic CGS unit of magnetic flux density. Deprecated.
giga	G	SI prefix for 10^9
gigaelectronvolt	GeV	
gigahertz	GHz	

TABLE 1-11 Standard Symbols for Units (*Continued*)

Unit	Symbol	Notes
gilbert	Gb	The gilbert is the electromagnetic CGS unit of magnetomotive force. Deprecated.
grain	gr	
gram	g	
gram per cubic centimeter	g/cm^3	
gray	Gy	SI unit of absorbed dose in the field of radiation dosimetry
hecto	h	SI prefix for 10^2
henry	H	SI unit of inductance
hertz	Hz	SI unit of frequency
horsepower	hp	The horsepower is an anachronism in science and technology. Use of the SI unit of power, the watt, is preferred.
hour	h	
inch	in	
conventional inch of mercury	inHg	1 inHg = 3386.4 Pa (ISO)
conventional inch of water	inH_2O	1 inH_2O = 249.09 Pa (ISO)
inch per second	in/s	
joule	J	SI unit of energy, work, quantity of heat
joule per kelvin	J/K	SI unit of heat capacity and entropy
kelvin	K	In 1967, the CGPM gave the name *kelvin* to the SI unit of temperature which had formerly been called *degree kelvin* and assigned it the symbol K (without the symbol °).
kilo	k	SI prefix for 10^3
kilogauss	kG	Deprecated.
kilogram	kg	SI unit of mass
kilogram-force	kgf	Deprecated. In some countries, the name kilopond (kp) has been used for this unit.
kilohertz	kHz	
kilohm	kΩ	
kilometer	km	
kilometer per hour	km/h	
kilopound-force	klbf	Kilopound-force should not be misinterpreted as kilopond (see kilogram-force).
kilovar	kvar	
kilovolt	kV	
kilovoltampere	kVA	
kilowatt	kW	
kilowatthour	kWh	Also kW · h
knot	kn	1kn = 1 nmi/h
lambert	L	1 L = $(1/\pi)$ cd/cm^2. A GGS unit of luminance. One lumen per square centimeter leaves a surface whose luminance is one lambert in all directions within a hemisphere. Deprecated.
liter	L	1 L = 10^{-3} m^3. The letter symbol 1 has been adopted for *liter* by the GGPM, and it is recommended in a number of international standards. In 1978, the CIPM accepted L as an alternative symbol. Because of frequent confusion with the numeral 1 the letter symbol 1 is no longer recommended for U.S. use. The script letter ℓ, which had been proposed, is not recommended as a symbol for liter.
liter per second	L/s	
lumen	lm	SI unit of luminous flux
lumen per square foot	lm/ft^2	A unit of illuminance and also a unit of luminous exitance. Use of the SI unit, lumen per square meter, is preferred.
lumen per square meter	lm/m^2	SI unit of luminous exitance
lumen per watt	lm/W	SI unit of luminous efficacy

(*Continued*)

TABLE 1-11 Standard Symbols for Units (*Continued*)

Unit	Symbol	Notes
lumen second	lm · s	SI unit of quantity of light
lux	lx	1 lx = 1 lm/m^2. SI unit of illuminance
maxwell	Mx	The maxwell is the electromagnetic CGS unit of magnetic flux. Deprecated.
mega	M	SI prefix for 10^6
megaelectronvolt	MeV	
megahertz	MHz	
megohm	MΩ	
meter	m	SI unit of length
metric ton	t	1 t = 1000 kg. The name *tonne* is used in some countries for this unit, but use of this name in the U.S. is deprecated.
mho	mho	Formerly used as the name of the siemens (S).
micro	μ	SI prefix for 10^{-6}
microampere	μA	
microfarad	μF	
microgram	μg	
microhenry	μH	
microinch	μin	
microliter	μL	See note for *liter*.
micrometer	μm	
micron	μm	Deprecated. Use micrometer.
microsecond	μs	
microwatt	μW	
mil	mil	1 mil = 0.001 in
mile (statute)	mi	1 mi = 5280 ft
miles per hour	mi/h	Although use of mph as an abbreviation is common, it should not be used as a symbol.
milli	m	SI prefix for 10^{-3}
milliampere	mA	
millibar	mbar	Use of the bar is strongly discouraged, except for limited use in meteorology.
milligram	mg	
millihenry	mH	
milliliter	mL	See note for *liter*.
millimeter	mm	
conventional millimeter of mercury	mmHg	1 mmHg = 133.322 Pa. Deprecated.
millimicron	nm	Use of the name *millimicron* for the nanometer is deprecated.
millipascal second	mPa · s	SI unit-multiple of dynamic viscosity
millisecond	ms	
millivolt	mV	
milliwatt	mW	
minute (plane angle)	. . . ′	
minute (time)	min	Time may also be designated by means of superscripts as in the following example: $9^h46^m30^s$.
mole	mol	SI unit of amount of substance
month	mo	
nano	n	SI prefix for 10^{-9}
nanoampere	nA	
nanofarad	nF	
nanometer	nm	
nanosecond	ns	
nautical mile	nmi	1 nmi = 1852 m

TABLE 1-11 Standard Symbols for Units (*Continued*)

Unit	Symbol	Notes
neper	Np	
newton	N	SI unit of force
newton meter	N · m	
newton per square meter	N/m^2	SI unit of pressure or stress, see pascal.
nit	nt	1 nt = 1 cd/m^2 The name *nit* is sometimes given to the SI unit of luminance, the candela per square meter.
oersted	Oe	The oersted is the electromagnetic CGS unit of magnetic field strength. Deprecated.
ohm	Ω	SI unit of resistance
ounce (avoirdupois)	oz	
pascal	Pa	1 Pa = 1 N/m^2 SI unit of pressure or stress
pascal second	Pa · s	SI unit of dynamic viscosity
peta	P	SI prefix for 10^{15}
phot	ph	1 ph = lm/cm^2 CGS unit of illuminance. Deprecated.
pico	p	SI prefix for 10^{-12}
picofarad	pF	
picowatt	pW	
pint	pt	1 pt (U.K.) = 0.568 26 L 1 pt (U.S. dry) = 0.550 61 L 1 pt (U.S. liquid) = 0.473 18 L
poise	P	Deprecated.
pound	lb	
pound per cubic foot	lb/ft^3	
pound-force	lbf	
pound-force foot	lbf · ft	
pound-force per square foot	lbf/ft^2	
pound-force per square inch	lbf/in^2	Although use of the abbreviation psi is common, it should not be used as a symbol.
poundal	pdl	
quart	qt	1 qt (U.K.) = 1.136 5 L 1 qt (U.S. dry) = 1.101 2 L 1 qt (U.S. liquid) = 0.946 35 L
rad	rd	A unit of absorbed dose in the field of radiation dosimetry. Use of the SI unit, the gray, is preferred. 1 rd = 0.01 Gy.
radian	rad	SI unit of plane angle
rem	rem	A unit of dose equivalent in the field of radiation dosimetry. Use of the SI unit, the sievert, is preferred. 1 rem = 0.01 Sv.
revolution per minute	r/min	Although use of rpm as an abbreviation is common, it should not be used as a symbol.
revolution per second	r/s	
roentgen	R	A unit of exposure in the field of radiation dosimetry
second (plane angle)	. . . ″	
second (time)	s	SI unit of time
siemens	S	1 S = 1 Ω^{-1} SI unit of conductance. The name mho has been used for this unit in the U.S.
sievert	Sv	SI unit of dose equivalent in the field of radiation dosimetry. Name adopted by the CIPM in 1978.
slug	slug	1 slug = 14.5939 kg
square foot	ft^2	
square inch	in^2	

(*Continued*)

TABLE 1-11 Standard Symbols for Units (*Continued*)

Unit	Symbol	Notes
square meter	m^2	
square meter per second	m^2/s	SI unit of kinematic viscosity
square millimeter per second	mm^2/s	SI unit-multiple of kinematic viscosity
square yard	yd^2	
steradian	sr	SI unit of solid angle
stilb	sb	1 sb = 1 cd/cm^2 A CGS unit of luminance. Deprecated.
stokes	St	Deprecated.
tera	T	SI prefix for 10^{12}
tesla	T	1 T = 1 N/(A · m) = 1 Wb/m^2. SI unit of magnetic flux density (magnetic induction).
therm	thm	1 thm = 100 000 Btu
ton (short)	ton	1 ton = 2000 lb
ton, metic	t	1 t = 1000 kg. The name *tonne* is used in some countries for this unit, but use of this name in the U.S. is deprecated.
(unified) atomic mass unit	u	The (unified) atomic mass unit is defined as one-twelfth of the mass of an atom of the ^{12}C nuclide. Use of the old atomic mass unit (amu), defined by reference to oxygen, is deprecated.
var	var	IEC name and symbol for the SI unit of reactive power
volt	V	SI unit of voltage
volt per meter	V/m	SI unit of electric field strength
voltampere	VA	IEC name and symbol for the SI unit of apparent power
watt	W	SI unit of power
watt per meter kelvin	W/(m · K)	SI unit of thermal conductivity
watt per steradian	W/sr	SI unit of radiant intensity
watt per steradian square meter	$W/(sr \cdot m^2)$	SI unit of radiance
watthour	Wh	
weber	Wb	Wb = V · s SI unit of magnetic flux
yard	yd	
year	a	In the English language, generally yr.

1.13 GRAPHIC SYMBOLS

An extensive list of standard graphic symbols for electrical engineering has been compiled in IEEE Standard 315 (ANSI Y32.2). Since this standard comprises 110 pages, including 78 pages of diagrams, it is impractical to reproduce it here. Those concerned with the preparation of circuit diagrams and graphic layouts should conform to these standard symbols to avoid confusion with earlier, nonstandard forms. See also Sec. 28.

1.14 PHYSICAL CONSTANTS

Table 1-12 lists the values of the fundamental physical constants, compiled by Peter, J. Mohr and Barry N. Taylor of the Task Group on Fundamental Constants of the Committee on Data for Science and Technology (CODATA), sponsored by the International Council of Scientific Unions. Further details on the methods used to adjust these values to form a consistent set are contained in Ref. 10. Table 1-13 lists the values of some energy equivalents.

TABLE 1-12 Fundamental Physical Universal Constants

Quantity	Symbol	Numerical value	Unit	Relative std. uncert. u_r
		UNIVERSAL		
speed of light in vacuum	c, c_0	299 792 458	$m\ s^{-1}$	(exact)
magnetic constant	μ_0	$4\pi \times 10^{-7}$	$N\ A^{-2}$	
		$= 12.566\ 370\ 614 \ldots \times 10^{-7}$	$N\ A^{-2}$	(exact)
electric constant $1/\mu_0 c^2$	ϵ_0	$8.854\ 187\ 817 \ldots \times 10^{-12}$	$F\ m^{-1}$	(exact)
characteristic impedance of vacuum $\sqrt{\mu_0/\epsilon_0} = \mu_0 c$	Z_0	376.730 313 461 …	Ω	(exact)
Newtonian constant of gravitation	G	$6.6742(10) \times 10^{-11}$	$m^3\ kg^{-1}\ s^{-2}$	1.5×10^{-4}
	$G/\hbar c$	$6.7087(10) \times 10^{-39}$	$(GeV/c^2)^{-2}$	1.5×10^{-4}
Planck constant	h	$6.626\ 0693(11) \times 10^{-34}$	J s	1.7×10^{-7}
in eV s		$4.135\ 667\ 43(35) \times 10^{-15}$	eV s	8.5×10^{-8}
$h/2\pi$	$\hbar$	$1.054\ 571\ 68(18) \times 10^{-34}$	J s	1.7×10^{-7}
in eV s		$6.582\ 119\ 15(56) \times 10^{-16}$	eV s	8.5×10^{-8}
$\hbar c$ in MeV fm		197.326 968(17)	Me V fm	8.5×10^{-8}
Planck mass $(\hbar c/G)^{1/2}$	m_P	$2.176\ 45(16) \times 10^{-8}$	kg	7.5×10^{-5}
Planck temperature $(\hbar c^5/G)^{1/2}/k$	T_P	$1.416\ 79(11) \times 10^{32}$	K	7.5×10^{-5}
Planck length $\hbar/m_P c = (\hbar G/c^3)^{1/2}$	l_P	$1.616\ 24(12) \times 10^{-35}$	m	7.5×10^{-5}
Planck time $l_P/c = (\hbar G/c^5)^{1/2}$	t_P	$5.391\ 21(40) \times 10^{-44}$	s	7.5×10^{-5}
		ELECTROMAGNETIC		
elementary charge	e	$1.602\ 176\ 53(14) \times 10^{-19}$	C	8.5×10^{-8}
	e/h	$2.417\ 989\ 40(21) \times 10^{14}$	$A\ J^{-1}$	8.5×10^{-8}
magnetic flux quantum h/2e	Φ_0	$2.067\ 833\ 72(18) \times 10^{-15}$	Wb	8.5×10^{-8}
conductance quantum $2e^2/h$	G_0	$7.748\ 091\ 733(26) \times 10^{-5}$	S	3.3×10^{-9}
inverse of conductance quantum	G_0^{-1}	12 906.403 725(43)	Ω	3.3×10^{-9}
Josephson constant $2e/h$	K_J	$483\ 597.879(41) \times 10^9$	$Hz\ V^{-1}$	8.5×10^{-8}
von Klitzing constant $h/e^2 = \mu_0 c/2\alpha$	R_K	25 812.807 449(86)	Ω	3.3×10^{-9}
Bohr magneton $e\hbar/2m_e$	μ_B	$927.400\ 949(80) \times 10^{-26}$	$J\ T^{-1}$	8.6×10^{-8}
in eV T^{-1}		$5.788\ 381\ 804(39) \times 10^{-5}$	$eV\ T^{-1}$	6.7×10^{-9}
	μ_B/h	$13.996\ 2458(12) \times 10^9$	$Hz\ T^{-1}$	8.6×10^{-8}
	μ_B/hc	46.686 4507(40)	$m^{-1}\ T^{-1}$	8.6×10^{-8}
	μ_B/k	0.671 7131(12)	$K\ T^{-1}$	1.8×10^{-6}
nuclear magneton $e\hbar/2m_P$	μ_N	$5.050\ 783\ 43(43) \times 10^{-27}$	$J\ T^{-1}$	8.6×10^{-8}
in eV T^{-1}		$3.152\ 451\ 259(21) \times 10^{-8}$	$eV\ T^{-1}$	6.7×10^{-9}
	μ_N/h	7.622 593 71(65)	$MHz\ T^{-1}$	8.6×10^{-8}
	μ_N/hc	$2.542\ 623\ 58(22) \times 10^{-2}$	$m^{-1}\ T^{-1}$	8.6×10^{-8}
	μ_N/k	$3.658\ 2637(64) \times 10^{-4}$	$K\ T^{-1}$	1.8×10^{-6}
		ATOMIC AND NUCLEAR		
		General		
fine-structure constant $e^2/4\pi\epsilon_0\hbar c$	α	$7.297\ 352\ 568(24) \times 10^{-3}$		3.3×10^{-9}
inverse fine-structure constant	α^{-1}	137.035 999 11(46)		3.3×10^{-9}
Rydberg constant $\alpha^2 m_e c/2h$	R_∞	10 973 731.568 525(73)	m^{-1}	6.6×10^{-12}
	$R_\infty c$	$3.289\ 841\ 960\ 360(22) \times 10^{15}$	Hz	6.6×10^{-12}
	$R_\infty hc$	$2.179\ 872\ 09(37) \times 10^{-18}$	J	1.7×10^{-7}
$R_\infty hc$ in eV		13.605 6923(12)	eV	8.5×10^{-8}
Bohr radius $\alpha/4\pi R_\infty = 4\pi\epsilon_0\hbar^2/m_e e^2$	a_0	$0.529\ 177\ 2108(18) \times 10^{-10}$	m	3.3×10^{-9}
Hartree energy $e^2/4\pi\epsilon_0 a_0 = 2R_\infty hc = \alpha^2 m_e c^2$	E_h	$4.359\ 744\ 17(75) \times 10^{-18}$	J	1.7×10^{-7}
in eV		27.211 3845(23)	eV	8.5×10^{-8}

(*Continued*)

TABLE 1-12 Fundamental Physical Universal Constants (*Continued*)

Quantity	Symbol	Numerical value	Unit	Relative std. uncert. u_r
quantum of circulation	$h/2m_e$	$3.636\,947\,550(24) \times 10^{-4}$	$m^2\ s^{-1}$	6.7×10^{-9}
	h/m_e	$7.273\,895\,101(48) \times 10^{-4}$	$m^2\ s^{-1}$	6.7×10^{-9}
	Electroweak			
Fermi coupling constant[a]	$G_F/(\hbar c)^3$	$1.166\,39(1) \times 10^{-5}$	GeV^{-2}	8.6×10^{-6}
weak mixing angle[b] θ_W (on-shell scheme)				
$\sin^2\theta_W = s^2_W \equiv 1 - (m\text{w}/mz)^2$	$\sin^2\theta_W$	0.222 15(76)		3.4×10^{-3}
	Electron, e^-			
electron mass	m_e	$9.109\,3826(16) \times 10^{-31}$	kg	1.7×10^{-7}
in u, $m_e = A_r(e)$ u (electron relative atomic mass times u)		$5.485\,799\,0945(24) \times 10^{-4}$	u	4.4×10^{-10}
energy equivalent	m_ec^2	$8.187\,1047(14) \times 10^{-14}$	J	1.7×10^{-7}
in MeV		0.510 998 918(44)	MeV	8.6×10^{-8}
electron-muon mass ratio	m_e/m_μ	$4.836\,331\,67(13) \times 10^{-3}$		2.6×10^{-8}
electron-tau mass ratio	m_e/m_τ	$2.875\,64(47) \times 10^{-4}$		1.6×10^{-4}
electron-proton mass ratio	m_e/m_p	$5.446\,170\,2173(25) \times 10^{-4}$		4.6×10^{-10}
electron-neutron mass ratio	m_e/m_n	$5.438\,673\,4481(38) \times 10^{-4}$		7.0×10^{-10}
electron-deuteron mass ratio	m_e/m_d	$2.724\,437\,1095(13) \times 10^{-4}$		4.8×10^{-10}
electron to alpha particle mass ratio	m_e/m_α	$1.370\,933\,555\,75(61) \times 10^{-4}$		4.4×10^{-10}
electron charge to mass quotient	$-e/m_e$	$-1.758\,820\,12(15) \times 10^{-11}$	$C\ kg^{-1}$	8.6×10^{-8}
electron molar mass N_Am_e	$M(e)$, M_e	$5.485\,799\,0945(24) \times 10^{-7}$	$kg\ mol^{-1}$	4.4×10^{-10}
Compton wavelength h/m_ec	λ_C	$2.426\,310\,238(16) \times 10^{-12}$	m	6.7×10^{-9}
$\lambda_C/2\pi = \alpha a_0 = \alpha^2/4\pi R_\infty$	$\bar{\lambda}_C$	$386.159\,2678(26) \times 10^{-15}$	m	6.7×10^{-9}
classical electron radius $\alpha^2 a_0$	r_e	$2.817\,940\,325(28) \times 10^{-15}$	m	1.0×10^{-8}
Thomson cross section $(8\pi/3)\,r_e^2$	σ_e	$0.665\,245\,873(13) \times 10^{-28}$	m^2	2.0×10^{-8}
electron magnetic moment	μ_e	$-928.476\,412(80) \times 10^{-26}$	$J\ T^{-1}$	8.6×10^{-8}
to Bohr magneton ratio	μ_e/μ_B	−1.001 159 652 1859(38)		3.8×10^{-12}
to nuclear magneton ratio	μ_e/μ_N	−1838.281 971 07(85)		4.6×10^{-10}
electron magnetic moment anomaly $\lvert\mu_e\rvert/\mu_B - 1$	a_e	$1.159\,652\,1859(38) \times 10^{-3}$		3.2×10^{-9}
electron *g*-factor $-2(1 + a_e)$	g_e	−2.002 319 304 3718(75)		3.8×10^{-12}
electron-muon magnetic moment ratio	μ_e/μ_μ	206.766 9894(54)		2.6×10^{-8}
electron-proton magnetic moment ratio	μ_e/μ_p	−658.210 6862(66)		1.0×10^{-8}
electron to shielded proton magnetic moment ratio (H_2O, sphere, 25 (C)	μ_e/μ'_p	−658.227 5956(71)		1.1×10^{-8}
electron-neutron magnetic moment ratio	μ_e/μ_n	960.920 50(23)		2.4×10^{-7}
electron-deuteron magnetic moment ratio	μ_e/μ_d	−2143.923 493(23)		1.1×10^{-8}
electron to shielded helion[c] magnetic moment ratio (gas, sphere, 25 °C)	μ_e/μ'_h	864.058 255(10)		1.2×10^{-8}
electron gyromagnetic ratio $2\lvert\mu_e\rvert/\hbar$	γ_e	$1.760\,859\,74(15) \times 10^{-11}$	$s^{-1}\ T^{-1}$	8.6×10^{-8}
	$\gamma_e/2\pi$	28 024.9532(24)	$MHz\ T^{-1}$	8.6×10^{-8}
		Muon, μ^-		
muon mass	m_μ	$1.883\,531\,40(33) \times 10^{-28}$	kg	1.7×10^{-7}
in u, $m_\mu = A_r(\mu)$ u (muon relative atomic mass time u)		0.113 428 9264(30)	u	2.6×10^{-8}
energy equivalent	$m_\mu c^2$	$1.692\,833\,60(29) \times 10^{-11}$	J	1.7×10^{-7}
in MeV		105.658 3692(94)	MeV	8.9×10^{-8}

TABLE 1-12 Fundamental Physical Universal Constants (*Continued*)

Quantity	Symbol	Numerical value	Unit	Relative std. uncert. u_r
muon-electron mass ratio	m_μ/m_e	206.768 2838(54)		2.6×10^{-8}
muon-tau mass ratio	m_μ/m_τ	$5.945\ 92(97) \times 10^{-2}$		1.6×10^{-4}
muon-proton mass ratio	m_μ/m_p	0.112 609 5269(29)		2.6×10^{-8}
muon-neutron mass ratio	m_μ/m_n	0.112 454 5175(29)		2.6×10^{-8}
muon molar mass $N_A m_\mu$	$M(\mu)$, M_μ	$0.113\ 428\ 9264(30) \times 10^{-3}$	kg mol^{-1}	2.6×10^{-8}
moun Compton wavelength $h/m_\mu c$	$\lambda_{C,\mu}$	$11.734\ 441\ 05(30) \times 10^{-15}$	m	2.5×10^{-8}
$\lambda_{C,\mu}/2\pi$	$\bar{\lambda}_{C,\mu}$	$1.867\ 594\ 298(47) \times 10^{-15}$	m	2.5×10^{-8}
moun magnetic moment	μ_μ	$-4.490\ 447\ 99(40) \times 10^{-26}$	J T^{-1}	8.9×10^{-8}
to Bohr magneton ratio	μ_μ/μ_B	$-4.841\ 970\ 45(13) \times 10^{-3}$		2.6×10^{-8}
to nuclear magneton ratio	μ_μ/μ_N	$-8.890\ 596\ 98(23)$		2.6×10^{-8}
muon magnetic moment anomaly $\lvert\mu_\mu\rvert/(e\hbar/2m_\mu) - 1$	a_μ	$1.165\ 919\ 81(62) \times 10^{-3}$		5.3×10^{-7}
moun g-factor $-2(1 + a_\mu)$	g_μ	$-2.002\ 331\ 8396(12)$		6.2×10^{-10}
moun-proton magnetic moment ratio	μ_μ/μ_p	$-3183\ 345\ 118(89)$		2.8×10^{-8}
	Tau, τ^-			
tau mass[d]	m_τ	$3.167\ 77(52) \times 10^{-27}$	kg	1.6×10^{-4}
in u, $m_\tau = A_r(\tau)$ u (tau relative atomic mass times u)		1.907 68(31)	u	1.6×10^{-4}
energy equivalent	$m_\tau c^2$	$2.847\ 05(46) \times 10^{-10}$	J	1.6×10^{-4}
in MeV		1776.99(29)	MeV	1.6×10^{-4}
tau-electron mass ratio	m_τ/m_e	3477.48(57)		1.6×10^{-4}
tau-muon mass ratio	m_τ/m_μ	16.8183(27)		1.6×10^{-4}
tau-proton mass ratio	m_τ/m_p	1.893 90(31)		1.6×10^{-4}
tau-neutron mass ratio	m_τ/m_n	1.891 29(31)		1.6×10^{-4}
tau molar mass $N_A m_\tau$	$M(\tau)$, M_τ	$1.907\ 68(31) \times 10^{-3}$	kg mol^{-1}	1.6×10^{-4}
tau Compton wavelength $h/m_\tau c$	$\lambda_{C,\tau}$	$0.697\ 72(11) \times 10^{-15}$	m	1.6×10^{-4}
$\lambda_{C,\tau}/2\pi$	$\bar{\lambda}_{C,\tau}$	$0.111\ 046(18) \times 10^{-15}$	m	1.6×10^{-4}
	Proton, p			
proton mass	m_p	$1.672\ 621\ 71(29) \times 10^{-27}$	kg	1.7×10^{-7}
in u, $m_p = A_r(p)$ u (proton relative atomic mass times u)		1.007 276 466 88(13)	u	1.3×10^{-10}
energy equivalent	$m_p c^2$	$1.503\ 277\ 43(26) \times 10^{-10}$	J	1.7×10^{-7}
in MeV		938.272 029(80)	MeV	8.6×10^{-8}
proton-electron mass ratio	m_p/m_e	1836.152 672 61(85)		4.6×10^{-10}
proton-muon mass ratio	m_p/m_μ	8.880 243 33(23)		2.6×10^{-8}
proton-tau mass ratio	m_p/m_τ	0.528 012(86)		1.6×10^{-4}
proton-neutron mass ratio	m_p/m_n	0.998 623 478 72(58)		5.8×10^{-10}
proton charge to mass quotient	e/m_p	$9.878\ 833\ 76(82) \times 10^{7}$	C kg^{-1}	8.6×10^{-8}
proton molar mass $N_A m_p$	$M(p)$, M_p	$1.007\ 276\ 466\ 88(13) \times 10^{-3}$	kg mol^{-1}	1.3×10^{-10}
proton Compton wavelength $h/m_p c$	$\lambda_{C,p}$	$1.321\ 409\ 8555(88) \times 10^{-15}$	m	6.7×10^{-9}
$\lambda_{C,p}/2\pi$	$\bar{\lambda}_{C,p}$	$0.210\ 308\ 9104(14) \times 10^{-15}$	m	6.7×10^{-9}
proton rms charge radius	R_p	$0.8750(68) \times 10^{-15}$	m	7.8×10^{-3}
proton magnetic moment	μ_p	$1.410\ 606\ 71(12) \times 10^{-26}$	J T^{-1}	8.7×10^{-8}
to Bohr magneton ratio	μ_p/μ_B	$1.521\ 032\ 206(15) \times 10^{-3}$		1.0×10^{-8}
to nuclear magneton ratio	μ_p/μ_N	2.792 847 351(28)		1.0×10^{-8}
proton g-factor $2\mu_p/\mu_N$	g_p	5.585 694 701(56)		1.0×10^{-8}
proton-neutron magnetic moment ratio	μ_p/μ_n	$-1.459\ 898\ 05(34)$		2.4×10^{-7}

(*Continued*)

TABLE 1-12 Fundamental Physical Universal Constants (*Continued*)

Quantity	Symbol	Numerical value	Unit	Relative std. uncert. u_r
shielded proton magnetic moment (H_2O, sphere, 25°C)	μ'_p	$1.410\ 570\ 47(12) \times 10^{-26}$	J T^{-1}	8.7×10^{-8}
to Bohr magneton ratio	μ'_p/μ_B	$1.520\ 993\ 132(16) \times 10^{-3}$		1.1×10^{-8}
to nuclear magneton ratio	μ'_p/μ_N	2.792 775 604(30)		1.1×10^{-8}
proton magnetic shielding correction $1 - \mu'_p/\mu_p$ (H_2O, sphere, 25°C)	σ'_p	$25.689(15) \times 10^{-6}$		5.7×10^{-4}
proton gyromagnetic ratio $2\mu_p/\hbar$	γ_p	$2.675\ 222\ 05(23) \times 10^{8}$	s^{-1} T^{-1}	8.6×10^{-8}
	$\gamma_p/2\pi$	42.577 4813(37)	MHz T^{-1}	8.6×10^{-8}
shielded proton gyromagnetic ratio $2\mu'_p/\hbar$ (H_2O, sphere, 25°C)	γ'_p	$2.675\ 153\ 33(23) \times 10^{8}$	s^{-1} T^{-1}	8.6×10^{-8}
	$\gamma'_p/2\pi$	42.576 3875(37)	MHz T^{-1}	8.6×10^{-8}
		Neutron, n		
neutron mass	m_n	$1.674\ 927\ 28(29) \times 10^{-27}$	kg	1.7×10^{-7}
in u, $m_n = A_r(n)$ u (neutron relative atomic mass times u)		1.008 664 915 60(55)	u	5.5×10^{-10}
energy equivalent	m_nc^2	$1.505\ 349\ 57(26) \times 10^{-10}$	J	1.7×10^{-7}
in MeV		939.565 360(81)	MeV	8.6×10^{-8}
neutron-electron mass ratio	m_n/m_e	1838.683 6598(13)		7.0×10^{-10}
neutron-muon mass ratio	m_n/m_μ	8.892 484 02(23)		2.6×10^{-8}
neutron-tau mass ratio	m_n/m_τ	0.528 740(86)		1.6×10^{-4}
neutron-proton mass ratio	m_n/m_p	1.001 378 418 70(58)		5.8×10^{-10}
neutron molar mass $N_A m_n$	M(n), M_n	$1.008\ 664\ 915\ 60(55) \times 10^{-3}$	kg mol^{-1}	5.5×10^{-10}
neutron Compton wavelength h/m_nc	$\lambda_{C,n}$	$1.319\ 590\ 9067(88) \times 10^{-15}$	m	6.7×10^{-9}
$\lambda_{C,n}/2\pi$	$\bar{\lambda}_{C,n}$	$0.210\ 019\ 4157(14) \times 10^{-15}$	m	6.7×10^{-9}
neutron magnetic moment	μ_n	$-0.966\ 236\ 45(24) \times 10^{-26}$	J T^{-1}	2.5×10^{-7}
to Bohr magneton ratio	μ_n/μ_B	$-1.041\ 875\ 63(25) \times 10^{-3}$		2.4×10^{-7}
to nuclear magneton ratio	μ_n/μ_N	−1.913 042 73(45)		2.4×10^{-7}
neutron g-factor $2\mu_n/\mu_N$	g_n	−3.826 085 46(90)		2.4×10^{-7}
neutron-electron magnetic moment ratio	μ_n/μ_e	$1.040\ 668\ 82(25) \times 10^{-3}$		2.4×10^{-7}
magnetic-proton magnetic moment ratio	μ_n/μ_p	−0.684 979 34(16)		2.4×10^{-7}
neutron to shielded proton magnetic moment ratio (H_2O, sphere, 25°C)	μ_n/μ'_p	−0.684 996 94(16)		2.4×10^{-7}
neutron gyromagnetic ratio $2\|\mu_n\|\hbar$	γ_n	$1.832\ 471\ 83(46) \times 10^{8}$	s^{-1} T^{-1}	2.5×10^{-7}
	$\gamma_n/2\pi$	29.164 6950(73)	MHz T^{-1}	2.5×10^{-7}
		Deuteron, d		
deuteron mass	m_d	$3.343\ 583\ 35(57) \times 10^{-27}$	kg	1.7×10^{-7}
in u, $m_d = A_r(d)$ u (deuteron relative atomic mass times u)		2.013 553 212 70(35)	u	1.7×10^{-10}
energy equivalent	m_dc^2	$3.005\ 062\ 85(51) \times 10^{-10}$	J	1.7×10^{-7}
in MeV		1875.612 82(16)	MeV	8.6×10^{-8}
deuteron-electron mass ratio	m_d/m_e	3670.482 9652(18)		4.8×10^{-10}
deuteron-proton mass ratio	m_d/m_p	1.999 007 500 82(41)		2.0×10^{-10}
deuteron molar mass $N_A\ m_d$	M(d), M_d	$2.013\ 553\ 212\ 70(35) \times 10^{-3}$	kg mol^{-1}	1.7×10^{-10}
deuteron rms charge radius	R_d	$2.1394(28) \times 10^{-15}$	m	1.3×10^{-3}

TABLE 1-12 Fundamental Physical Universal Constants (*Continued*)

Quantity	Symbol	Numerical value	Unit	Relative std. uncert. u_r
deuteron magnetic moment	μ_d	$0.433\ 073\ 482(38) \times 10^{-26}$	J T^{-1}	8.7×10^{-8}
to Bohr magneton ratio	μ_d/μ_B	$0.466\ 975\ 4567(50) \times 10^{-3}$		1.1×10^{-8}
to nuclear magneton ratio	μ_d/μ_N	0.857 438 2329(92)		1.1×10^{-8}
deuteron-electron magnetic moment ratio	μ_d/μ_e	$-4.664\ 345\ 548(50) \times 10^{-4}$		1.1×10^{-8}
deuteron-proton magnetic moment ratio	μ_d/μ_p	0.307 012 2084(45)		1.5×10^{-8}
deuteron-neutron magnetic moment ratio	μ_d/μ_n	–0.448 206 52(11)		2.4×10^{-7}
		Helion, h		
helion mass[c]	m_h	$5.006\ 412\ 14(86) \times 10^{-27}$	kg	1.7×10^{-7}
in u, $m_h = A_r$(h) u (helion relative atomic mass times u)		3.014 932 2434(58)	u	1.9×10^{-9}
energy equivalent	$m_h c^2$	$4.499\ 538\ 84(77) \times 10^{-10}$	J	1.7×10^{-7}
in MeV		2808.391 42(24)	MeV	8.6×10^{-8}
helion-electron mass ratio	m_h/m_e	5495.885 269(11)		2.0×10^{-9}
helion-proton mass ratio	m_h/m_p	2.993 152 6671(58)		1.9×10^{-9}
helion molar mass $N_A m_h$	M(h), M_h	$3.014\ 932\ 2434(58) \times 10^{-3}$	kg mol^{-1}	1.9×10^{-9}
shielded helion magnetic moment (gas, sphere, 25°C)	μ'_h	$-1.074\ 553\ 024(93) \times 10^{-26}$	J T^{-1}	8.7×10^{-8}
to Bohr magneton ratio	μ'_h/μ_B	$-1.158\ 671\ 474(14) \times 10^{-3}$		12×10^{-8}
to nuclear magneton ratio	μ'_h/μ_N	–2.127 497 723(25)		12×10^{-8}
shielded helion to proton magnetic moment ratio (gas, sphere, 25°C)	μ'_h/μ_p	–0.761 766 562(12)		1.5×10^{-8}
shielded helion to shielded proton magnetic moment ratio (gas/H_2O, spheres, 25°C)	μ'_h/μ'_p	–0.761 786 1313(33)		4.3×10^{-9}
shielded helion gyromagnetic ratio $2\|\mu'_h\|/\hbar$ (gas, sphere, 25°C)	γ'_h	$2.037\ 894\ 70(18) \times 10^{8}$	s^{-1} T^{-1}	8.7×10^{-8}
	$\gamma'_h/2\pi$	32.434 1015(28)	MHz T^{-1}	8.7×10^{-8}
		Alpha particle, α		
alpha particle mass	m_α	$6.644\ 6565(11) \times 10^{-27}$	kg	1.7×10^{-7}
in u, $m_\alpha = A_r(\alpha)$ u (alpha particle relative atomic mass times u)		4.001 506 179 149(56)	u	1.4×10^{-11}
energy equivalent	$m_\alpha c^2$	$5.971\ 9194(10) \times 10^{-10}$	J	1.7×10^{-7}
in MeV		3727.379 17(32)	MeV	8.6×10^{-8}
alpha particle to electron mass ratio	m_α/m_e	7294.299 5363(32)		4.4×10^{-10}
alpha particle to proton mass ratio	m_α/m_p	3.972 599 689 07 (52)		1.3×10^{-10}
alpha particle molar mass $N_A m_\alpha$	$M(\alpha)$, M_α	$4.001\ 506\ 179\ 149(56) \times 10^{-3}$	kg mol^{-1}	1.4×10^{-11}
		PHYSICO-CHEMICAL		
Avogadro constant	N_A, L	$6.022\ 1415(10) \times 10^{23}$	mol^{-1}	1.7×10^{-7}
atomic mass constant $m_u = {}^1/_{12}m(^{12}C) = 1$ u $= 10^{-3}$ kg mol^{-1}/N_A	m_u	$1.660\ 538\ 86(28) \times 10^{-27}$	kg	1.7×10^{-7}
energy equivalent	$m_u c^2$	$1.492\ 417\ 90(26) \times 10^{-10}$	J	1.7×10^{-7}
in MeV		931.494 043(80)	MeV	8.6×10^{-8}
Faraday constant[e] $N_A e$	F	96 485.3383(83)	C mol^{-1}	8.6×10^{-8}
molar Planck constant	$N_A h$	$3.990\ 312\ 716(27) \times 10^{-10}$	J s mol^{-1}	6.7×10^{-9}
	$N_A hc$	0.119 626 565 72(80)	J m mol^{-1}	6.7×10^{-9}

(*Continued*)

TABLE 1-12 Fundamental Physical Universal Constants (*Continued*)

Quantity	Symbol	Numerical value	Unit	Relative std. uncert. u_r
molar gas constant	R	8.314 472(15)	J mol^{-1} K^{-1}	1.7×10^{-6}
Boltzmann constant R/N_A	k	$1.380\ 6505(24) \times 10^{-23}$	J K^{-1}	1.8×10^{-6}
in eV K^{-1}		$8.617\ 343(15) \times 10^{-5}$	eV K^{-1}	1.8×10^{-6}
	k/h	$2.083\ 6644(36) \times 10^{10}$	Hz K^{-1}	1.7×10^{-6}
	k/hc	69.503 56(12)	m^{-1} K^{-1}	1.7×10^{-6}
molar volume of ideal gas RT/p				
T = 273.15 K, p = 101.325 kpa	V_m	$22.413\ 996(39) \times 10^{-3}$	m^3 mol^{-1}	1.7×10^{-6}
Loschmidt constant N_A/V_m	n_0	$2.686\ 7773(47) \times 10^{25}$	m^{-3}	1.8×10^{-6}
T = 273.15 K, p = 100 kpa	V_m	$22.710\ 981(40) \times 10^{-3}$	m^3 mol^{-1}	1.7×10^{-6}
Sackur-Tetrode constant (absolute entropy constant)[f]				
$5/2 + \ln[2\pi m_u kT_1/h^2)^{3/2} kT_1/p_0]$				
T_1 = 1 K, p_0 = 100 kPa	S_0/R	−1.151 7047(44)		3.8×10^{-6}
T_1 = 1 K, p_0 = 101.325 kPa		−1.164 8677(44)		3.8×10^{-6}
Stefan-Boltzmann constant				
$(\pi^2/60)\ k^4/\hbar^3 c^2$	σ	$5.670\ 400(40) \times 10^{-8}$	W m^{-2} K^{-4}	7.0×10^{-6}
first radiation constant $2\pi hc^2$	c_1	$3.741\ 771\ 38(64) \times 10^{-16}$	W m^2	1.7×10^{-7}
first radiation constant for spectral radiance $2hc^2$	c_{1L}	$1.191\ 042\ 82(20) \times 10^{-16}$	W m^2 sr^{-1}	1.7×10^{-7}
second radiation constant hc/k	c_2	$1.438\ 7752(25) \times 10^{-2}$	m K	1.7×10^{-6}
Wien displacement law constant				
$b = \lambda_{max}T = c_2/4.965\ 114\ 231\ldots$	b	$2.897\ 7685(51) \times 10^{-3}$	m K	1.7×10^{-6}

Source: *CODATA recommended values of the fundamental physical constants: 2002; Peter J. Mohr and Barry N. Taylor; *Rev, Mod, Phys.* January 2005, vol. 77, no. 1, pp. 1–107.

[a] Value recommended by the Particle Data Group (Hagiwara et al., 2002).

[b] Based on the ratio of the masses of the W and Z bosons m_W/m_Z recommended by the Particle Data Group (Hagiwara *et al.*, 2002). The value for $\sin^2 \theta_W$ they recommend, which is based on a particular variant of the modified minimal subtraction ($\overline{\text{MS}}$) scheme, is $\sin^2 \hat{\theta}_W (Mz) = 0.231\ 24(24)$.

[C] The hellion, symbol h, is the nucleus of the ^{3}He atom.

[d] This and all other values involving m_τ are based on the value of $m_\tau c^2$ in MeV recommended by the Particle Data Group (Hagiwara *et al.*, 2002), but with a standard uncertainty of 0.29 MeV rather than the quoted uncertainty of –0.26 MeV, +0.29 MeV.

[e] The numerical value of F to be used in coulometric chemical measurements is 96 485.336(16) [1.7×10^{-7}] when the relevant current is measured in terms of representations of the volt and ohm based on the Josephson and quantum Hall effects and the internationally, adopted conventional values of the Josephson and von Klitzing constants K_{J-90} and R_{K-90}.

[f] The entropy of an ideal monoatomic gas of relative atomic mass A_r is given by $S = S_0 + 3/2\ R \ln A_r - R \ln (p/p_0) + 5/2\ R \ln (T/K)$.

1.15 NUMERICAL VALUES

Extensive use is made in electrical engineering of the constants π and ϵ and of the numbers 2 and 10, the latter in logarithmic units and number systems. Table 1-14 lists functions of these numbers to 9 or 10 significant digits. In most engineering applications (except those involving the difference of large, nearly equal numbers), five significant digits suffice. The use of the listed values in computations with electronic hand calculators will suffice in most cases to produce results more than adequate for engineering work.

1.16 CONVERSION FACTORS

The increasing use of the metric system in British and American practice has generated a need for extensive tables of multiplying factors to facilitate conversions from and to the SI units. Tables 1-15 through 1-28 list these conversion factors.

Table	Quantity	SI unit	Subtabulation	Basis of grouping
1-15	Length	meter	1-15A	Units decimally related to one meter
			1-15B	Units less than one meter
			1-15C	Units greater than one meter
			1-15D	Other length units
1-16	Area	square meter	1-16A	Units decimally related to one square meter
			1-16B	Nonmetric area units
			1-16C	Other area units
1-17	Volume/capacity	cubic meter	1-17A	Units decimally related to one cubic meter
			1-17B	Nonmetric volume units
			1-17C	U.S. liquid capacity measures
			1-17D	British liquid capacity measures
			1-17E	U.S. and U.K. dry capacity measures
			1-17F	Other volume and capacity units
1-18	Mass	kilogram	1-18A	Units decimally related to one kilogram
			1-18B	Less than one pound-mass
			1-18C	One pound-mass and greater
			1-18D	Other mass units
1-19	Time	second	1-19A	One second and less
			1-19B	One second and greater
			1-19C	Other time units
1-20	Velocity	meter per second		
1-21	Density	kilogram per cubic meter	1-21A	Units decimally related to one kilogram per cubic meter
			1-21B	Nonmetric density units
			1-21C	Other density units
1-22	Force	newton		
1-23	Pressure	pascal	1-23A	Units decimally related to one pascal
			1-23B	Units decimally related to one kilogram-force per square meter
			1-23C	Units expressed as heights of liquid
			1-23D	Nonmetric pressure units
1-24	Torque/bending moment	newton meter		
1-25	Energy/work	joule	1-25A	Units decimally related to one joule
			1-25B	Units less than 10 joules
			1-25C	Units greater than 10 joules
1-26	Power	watt	1-26A	Units decimally related to one watt
			1-26B	Nonmetric power units
1-27	Temperature	kelvin		
1-28	Light	candela per square meter	1-28A	Luminance units
		lux	1-28B	Illuminance units

Statements of Equivalence. To avoid ambiguity, the conversion tables have been arranged in the form of statements of equivalence, that is, each unit listed at the left-hand edge of each table is stated to be equivalent to a multiple or fraction of each of the units to the right in the table. For example, the uppermost line of Table 1-15B represents the following statements:

Column 2. 1 meter is equal to 1.093 613 30 yards

Column 3. 1 meter is equal to 3.280 839 89 feet

Column 4. 1 meter is equal to 39.370 078 7 inches

Column 5. 1 meter is equal to $3.937\ 007\ 87 \times 10^4$ mils

Column 6. 1 meter is equal to $3.937\ 007\ 87 \times 10^7$ microinches

TABLE 1-13 Derived Energy Equivalents

[Derived from the relations $E = mc^2 = hc/\lambda = hv = kT$, and based on the 2002 CODATA adjustment of the values of the constants; 1 eV = (e/C) J, 1 u = $m_u = \frac{1}{12} m(^{12}\mathrm{C}) = 10^{-3}$ kg mol$^{-1}/N_A$, and $E_h = 2R_\infty hc = \alpha^2 m_e c^2$ is the Hartree energy (hartree).]

	Relevant unit			
	J	kg	m^{-1}	Hz
1 J	(1 J) = 1 J	(1 J)/c^2 = 1.112 650 056… × 10^{-17} kg	(1 J)/hc = 5.034 117 20(86) × 10^{24} m^{-1}	(1 J)/h = 1.509 190 37(26) × 10^{33} Hz
1 kg	(1 kg)c^2 = 8.987 551 787… × 10^{16} J	(1 kg) = 1 kg	(1 kg) c/h = 4.524 438 91(77) × 10^{41} m^{-1}	(1 kg) c^2/h = 1.356 392 66(23) × 10^{50} Hz
1 m^{-1}	(1 m^{-1}) hc = 1.986 445 61(34) × 10^{-25} J	(1 m^{-1}) h/c = 2.210 218 81(38) × 10^{-42} kg	(1 m^{-1}) = 1m^{-1}	(1 m^{-1}) c = 299 792 458 Hz
1 Hz	(1 Hz) h = 6.626 0693(11) × 10^{-34} J	(1 Hz) h/c^2 = 7.372 4964(13) × 10^{-51} kg	(1 Hz)/c = 3.335 640 951… × 10^{-9} m^{-1}	(1 Hz) = 1 Hz
1 K	(1 K) k = 1.380 6505(24) × 10^{-23} J	(1 K) k/c^2 = 1.536 1808(27) × 10^{-40} kg	(1 K)k/hc = 69.503 56(12) m^{-1}	(1 K) k/h = 2.083 6644(36) × 10^{10} Hz
1 eV	(1 eV) = 1.602 176 53(14) × 10^{-19} J	(1 eV) /c^2 = 1.782 661 81(15) × 10^{-36} kg	(1 eV)/hc = 8.065 544 45 (69) × 10^{5} m^{-1}	(1 eV)/h = 2.417 989 40(21) × 10^{14} Hz
1 u	(1 u)c^2 = 1.492 417 90(26) × 10^{-10} J	(1 u) = 1.660 538 86(28) × 10^{-27} kg	(1 u)c/h = 7.513 006 608(50) × 10^{14} m^{-1}	(1 u) c^2/h = 2.252 342 718(15) × 10^{23} Hz
1 E_h	(1 E_h) = 4.359 744 17(75) × 10^{-18} J	(1 E_h)/c^2 = 4.850 869 60 (83) × 10^{-35} kg	(1 E_h)/hc = 2.194 746 313 705(15) × 10^{7} m^{-1}	(1 E_h)/h = 6.579 683 920 721(44) × 10^{15} Hz

	Relevant unit			
	K	eV	u	E_h
1 J	(1 J)/k = 7.242 963(13) × 10^{22} K	(1 J) = 6.241 509 47(53) ×10^{18} eV	(1 J)/c^2 = 6.700 5361(11) × 10^{9} u	(1 J) = 2.293 712 57(39) × 10^{17} E_h
1 kg	(1 kg)c^2/k = 6.509 650(11) × 10^{39} K	(1 kg)c^2 5.609 588 96(48) × 10 ^{35} eV	(1 kg)= 6.022 1415(10) × 10^{26} u	(1 kg)c^2 = 2.061 486 05(35) × 10^{34} E_h
1 m^{-1}	(1 m^{-1})hc/k = 1.438 7752(25) × 10^{-2} K	(1 m^{-1})hc = 1.239 841 91(11) × 10^{-6} eV	(1 m^{-1})h/c = 1.331 025 0506(89) × 10^{-15} u	(1 m^{-1})hc = 4.556 335 252 760(30) × 10^{-8} E_h
1 Hz	(1 Hz)h/k = 4.799 2374(84) × 10^{-11} K	(1 Hz)h = 4.135 667 43(35) × 10^{-15} eV	(1 Hz)h/c^2 = 4.439 821 667(30) × 10^{-24} u	(1 Hz)h = 1.519 829 846 006(10) × 10^{-16} E_h
1 K	(1 K) = 1 K	(1 K)k = 8.617 343(15) ×10^{-5} eV	(1 K)k/c^2 = 9.251 098(16) × 10^{-14} u	(1 K)k = 3.166 8153(55) × 10^{-6} E_h
1 eV	(1 eV)/k = 1.160 4505(20) × 10^{4} K	(1 eV) = 1 eV	(1 eV)/c^2 = 1.073 544 171(92) × 10^{-9} u	(1 eV) = 3.674 932 45(31) × 10^{-2} E_h
1 u	(1 u)c^2/k = 1.080 9527(19) × 10^{13} K	(1 u)c^2 = 931.494 043(80) × 10^{6} eV	(1 u)= 1 u	(1 u)c^2 = 3.423 177 686(23) × 10^{7} E_h
1 E_h	(1 E_h)/k = 3.157 7465(55) × 10^{5} K	(1 E_h) = 27.211 3845(23) eV	(1 E_h)/c^2 = 2.921 262 323(19) × 10^{-8} u	(1 E_h) = 1 E_h

TABLE 1-14 Numerical Values Used in Electrical Engineering

Functions of π:

$\pi = 3.141\ 592\ 654$

$1/\pi = 0.318\ 309\ 886$

$\pi^2 = 9.869\ 604\ 404$

$\sqrt{\pi} = 1.772\ 453\ 851$

$\pi/180° = 0.017\ 453\ 293$ (= radians per degree)

$180°/\pi = 57.295\ 779\ 51$ (= degrees per radian)

Functions of ϵ:

$\epsilon = 2.718\ 281\ 828$

$1/\epsilon = 0.367\ 879\ 441$

$1 - 1/\epsilon = 0.632\ 120\ 559$

$\epsilon^2 = 7.389\ 056\ 096$

$\sqrt{\epsilon} = 1.648\ 721\ 271$

(*Continued*)

TABLE 1-14 Numerical Values Used in Electrical Engineering (*Continued*)

Logarithms to the base 10:

$\log_{10} \pi = 0.497\ 149\ 873$

$\log_{10} \epsilon = 0.434\ 294\ 482$

$\log_{10} 2 = 0.301\ 029\ 996$

$\log_{10} x = (\ln x)(0.434\ 294\ 482) = (\log_2 x)(0.301\ 029\ 996)$

Natural logarithms (to the base ϵ):

$\ln \pi = 1.144\ 729\ 886$

$\ln 2 = 0.693\ 147\ 181$

$\ln 10 = 2.302\ 585\ 093$

$\ln x = (\log_{10} x)(2.302\ 585\ 093) = (\log_2 x)(0.693\ 147\ 181)$

Logarithms to the base 2:

$\log_2 \pi = 1.651\ 496\ 130$

$\log_2 \epsilon = 1.442\ 695\ 042$

$\log_2 10 = 3.321\ 928\ 096$

$\log_2 x = (\log_{10} x)(3.321\ 928\ 096) = (\ln x)(1.442\ 695\ 042)$

Powers of 2:

$2^5 = 32$

$2^{10} = 1024$

$2^{15} = 32{,}768$

$2^{20} = 1{,}048{,}576$

$2^{25} = 33{,}554{,}432$

$2^{30} = 1{,}073{,}741{,}824$

$2^{40} = 1.099\ 511\ 628 \times 10^{12}$

$2^{50} = 1.125\ 899\ 907 \times 10^{15}$

$2^{100} = 1.267\ 650\ 601 \times 10^{30}$

Logarithmic units:

Power ratio	Current or voltage ratio	Decibels*	Nepers†
1	1	0	0
2	1.414 214	3.010 300	0.346 574
3	1.732 051	4.771 213	0.549 306
4	2	6.020 600	0.693 147
5	2.236 068	6.989 700	0.804 719
10	3.162 278	10	1.151 293
15	3.872 983	11.760 913	1.354 025

Values of $2^{(2N)}$:

Value of N	Value of $2^{(2N)}$
1	4
2	16
3	256
4	65,536
5	4,294,967,296
6	$1.844\ 674\ 407 \times 10^{19}$
7	$3.402\ 823\ 668 \times 10^{38}$
8	$1.157\ 920\ 892 \times 10^{77}$
9	$1.340\ 780\ 792 \times 10^{154}$
10	$1.797\ 693\ 132 \times 10^{308}$

*The decibel is defined for power ratios only. It may be applied to current or voltage ratios only when the resistances through which the currents flow or across which the voltages are applied are equal.

†The neper is defined for current and voltage ratios only. It may be applied to power ratios only when the respective resistances are equal.

TABLE 1-15 Length Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of length is the meter.

A. Length units decimally related to one meter

	Meters (m)	Kilometers (km)	Decimeters (dm)	Centimeters (cm)	Millimeters (mm)	Micrometers (μm)	Nanometers (nm)	Ångströms (Å)
1 meter =	**1**	**0.001**	**10**	**100**	**1 000**	**1 000 000**	**10^{9}**	**10^{10}**
1 kilometer =	**1 000**	**1**	**10 000**	**100 000**	**1 000 000**	**10^{9}**	**10^{12}**	**10^{13}**
1 decimeter =	**0.1**	**0.000 1**	**1**	**10**	**100**	**100 000**	**10^{8}**	**10^{8}**
1 centimeter =	**0.01**	**0.000 01**	**0.1**	**1**	**10**	**10 000**	**10^{7}**	**10^{8}**
1 millimeter =	**0.001**	**10^{-6}**	**0.01**	**0.1**	**1**	**1 000**	**1 000 000**	**10^{7}**
1 micrometer (micron) =	**10^{-6}**	**10^{-9}**	**0.000 01**	**0.000 1**	**0.001**	**1**	**1 000**	**10 000**
1 nanometer =	**10^{-9}**	**10^{-12}**	**10^{-8}**	**10^{-7}**	**10^{-6}**	**0.001**	**1**	**10**
1 ångström =	**10^{-10}**	**10^{-13}**	**10^{-9}**	**10^{-8}**	**10^{-7}**	**0.000 1**	**0.1**	**1**

B. Nonmetric length units less than one meter

	Meters (m)	Yards (yd)	Feet (ft)	Inches (in)	Mils (mil)	Microinches (μin)
1 meter =	**1**	1.093 613 30	3.280 839 89	39.370 078 7	$3.937\ 007\ 87 \times 10^{4}$	$3.937\ 007\ 87 \times 10^{7}$
1 yard =	**0.914 4**	**1**	**3**	**36**	**36 000**	**3.6×10^{7}**
1 foot =	**0.304 8**	**1/3** = $0.333\ \underline{3}$	**1**	**12**	**12 000**	**1.2×10^{7}**
1 inch =	**0.025 4**	**1/36** = $0.027\ \underline{7}$	**1/12** = $0.083\ \underline{3}$	**1**	**1 000**	**1 000 000**
1 mil =	**2.54×10^{-5}**	$2.77\underline{7} \times 10^{-5}$	$8.33\underline{3} \times 10^{-5}$	**0.001**	**1**	**1 000**
1 microinch =	**2.54×10^{-8}**	$2.77\underline{7} \times 10^{-8}$	$8.33\underline{3} \times 10^{-8}$	**10^{-8}**	**0.001**	**1**

C. Nonmetric length units greater than one meter (with equivalents in feet)

	Meters (m)	Rods (rd)	Statute miles (mi)	Nautical miles (nmi)	Astronomical units (AU)	Parsecs (pc)	Feet (ft)
1 meter =	**1**	0.198 838 78	$6.213\ 711\ 92 \times 10^{-4}$	$5.399\ 568\ 04 \times 10^{-4}$	$6.684\ 491\ 98 \times 10^{-12}$	$3.240\ 733\ 17 \times 10^{-17}$	3.280 839 89
1 rod =	**5.029 2**	**1**	**0.003 125**	$2.715\ 550\ 76 \times 10^{-3}$	$3.361\ 764\ 71 \times 10^{-11}$	$1.629\ 829\ 53 \times 10^{-16}$	**16.5**
1 statute mile =	**1 609.344**	**320**	**1**	0.868 976 24	$1.075\ 764\ 71 \times 10^{-8}$	$5.215\ 454\ 50 \times 10^{-14}$	**5 280**
1 nautical mile =	**1 852**	368.249 423	1.150 779 45	**1**	$1.237\ 967\ 91 \times 10^{-8}$	$6.001\ 837\ 80 \times 10^{-14}$	6 076.115 48
1 astronomical unit* =	**1.496×10^{11}**	$2.974\ 628\ 17 \times 10^{10}$	92 957 130.3	80 777 537.8	**1**	$4.848\ 136\ 82 \times 10^{-6}$	$4.908\ 136\ 48 \times 10^{11}$
1 parsec =	$3.085\ 721\ 50 \times 10^{16}$	$6.135\ 611\ 02 \times 10^{15}$	$1.917\ 378\ 44 \times 10^{13}$	$1.666\ 156\ 32 \times 10^{13}$	206 264.806	**1**	$1.012\ 375\ 82 \times 10^{17}$
1 foot =	**0.304 8**	$0.060\ \underline{606}$	$1.893\ 9\underline{39} \times 10^{-4}$	$1.645\ 788\ 33 \times 10^{-4}$	$2.037\ 433\ 16 \times {}^{-12}$	$9.877\ 754\ 72 \times 10^{-18}$	**1**

D. Other length units
1 cable = **720** feet = **219.456** meters
1 cable (U.K.) = **608** feet = **185.318 4** meters
1 chain (engineers') = **100** feet = **30.48** meters
1 chain (surveyors') = **66** feet = **20.116 8** meters
1 fathom = **6** feet = **1.828 8** meters
1 fermi = **1** femtometer = $\mathbf{10^{-15}}$ meter
1 foot (U.S. Survey) = 0.304 800 6 meter
1 furlong = **660** feet = **201.168** meters
1 hand = **4** inches = **0.101 6** meter
1 league (international nautical) = 3 nautical miles = **5 556** meters
1 league (statute) = **3** statute miles = **4 828.032** meters
1 league (U.K. nautical) = 5 559.552 meters
1 light-year = $9.460\,895\,2 \times 10^{15}$ meters (= distance traveled by light in vacuum in one sidereal year)
1 link (engineers') = **1** foot = **0.304 8** meter
1 link (surveyors') = **7.92** inches = **0.201 168** meter
1 micron = **1** micrometer = $\mathbf{10^{-6}}$ meter
1 millimicron = **1** nanometer = $\mathbf{10^{-9}}$ meter
1 myriameter = **10 000** meters
1 nautical mile (U.K.) = 1 853.184 meters
1 pale = **1** rod = **5.029 2** meters
1 perch (linear) = **1** rod = **5.029 2** meters
1 pica = **1/6** inch (approx.) = $4.217\,518 \times 10^{-3}$ meter
1 point = **1/72** inch (approx.) = $3.514\,598 \times 10^{-4}$ meter
1 span = **9** inches = **0.228 6** meter

*As defined by the International Astronomical Union.

TABLE 1-16 Area Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of area is the square meter.

A. Area units decimally related to one square meter

	Square meters $(m)^2$	Square kilometers $(km)^2$	Hectares **(square hectometers) $(hm)^2$**	Square centimeters $(cm)^2$	Square millimeters $(mm)^2$	Square micrometers $(\mu m)^2$	Barns (b)
1 square meter =	**1**	**10^{-6}**	**0.000 1**	**10 000**	**1 000 000**	**10^{12}**	**10^{28}**
1 square kilometer =	**1 000 000**	**1**	**100**	**10^{10}**	**10^{12}**	**10^{18}**	**10^{34}**
1 hectare =	**10 000**	**0.01**	**1**	**10^{8}**	**10^{10}**	**10^{16}**	**10^{32}**
1 square centimeter =	**0.000 1**	**10^{-10}**	**10^{-8}**	**1**	**100**	**10^{8}**	**10^{24}**
1 square millimeter =	**10^{-6}**	**10^{-12}**	**10^{-10}**	**0.01**	**1**	**10^{6}**	**10^{22}**
1 square micrometer =	**10^{-12}**	**10^{-18}**	**10^{-16}**	**10^{-8}**	**10^{-6}**	**1**	**10^{16}**
1 barn =	**10^{-28}**	**10^{-34}**	**10^{-32}**	**10^{-24}**	**10^{-22}**	**10^{-16}**	**1**

B. Nonmetric area units (with square meter equivalents)

	Square meters $(m)^2$	Square statute miles $(mi)^2$	Acres (acre)	Square rods $(rd)^2$	Square yards $(yd)^2$	Square feet $(ft)^2$	Square inches $(in)^2$	Circular mils (cmil)
1 square meter =	**1**	$3.861\,021\,59 \times 10^{-7}$	$2.471\,053\,82 \times 10^{-4}$	$3.953\,686\,10 \times 10^{-2}$	1.195 990 05	10.763 910 4	1 550.003 10	$1.973\,525\,24 \times 10^{9}$
1 square statute mile =	2 589 988.1	**1**	**640**	**102 400**	**3 097 600**	**27 878 400**	$4.014\,489\,60 \times 10^{9}$	$5.111\,406\,91 \times 10^{15}$
1 acre =	4 046.856 11	**1/640 = 0.001 562 5**	**1**	**160**	**4 840**	**43 560**	**6 272 640**	$7.986\,573\,30 \times 10^{12}$
1 square rod =	25.292 852 6	**$9.765\,625 \times 10^{-6}$**	**1/160 − 0.006 25**	**1**	**30.25**	**272.25**	**39 204**	$4.991\,608\,31 \times 10^{10}$
1 square yard =	**0.836 127 36**	$3.228\,305\,79 \times 10^{-7}$	$2.066\,115\,70 \times 10^{-4}$	$3.305\,785\,12 \times 10^{-2}$	**1**	**9**	**1 296**	$1.650\,118\,45 \times 10^{9}$
1 square foot =	**0.092 903 04**	$3.587\,006\,43 \times 10^{-8}$	$2.295\,684\,11 \times 10^{-5}$	$3.673\,094\,58 \times 10^{-3}$	**1/9** = 0.111 111	**1**	**144**	$1.833\,464\,95 \times 10^{8}$
1 square inch =	**$6.451\,6 \times 10^{-4}$**	$2.490\,976\,69 \times 10^{-10}$	$1.594\,225\,08 \times 10^{-7}$	$2.550\,760\,13 \times 10^{-5}$	$7.716\,049\,38 \times 10^{-4}$	**1/144** = 0.006 944 44	**1**	$1.273\,239\,55 \times 10^{6}$
1 circular mil =	$5.067\,074\,79 \times 10^{-10}$	$1.956\,408\,51 \times 10^{-16}$	$1.252\,101\,45 \times 10^{-13}$	$2.003\,362\,32 \times 10^{-11}$	$6.060\,171\,01 \times 10^{-10}$	$5.454\,153\,91 \times 10^{-9}$	$7.853\,981\,63 \times 10^{-7}$	**1**

Exact conversions are:

1 acre = **4 046.856 422 4** square meters

1 square mile = **2 589 988.110 336** square meters

C. Other area units

1 are = **100** square meters

1 centiare (centare) = **1** square meter

1 perch (area) = **1** square rod = **30.25** square yards = 25.292 852 6 square meters

1 rod = **40** square rods = 1 011.714 11 square meters

1 section = **1** square statute mile = 2 589 988.1 square meters

1 township = **36** square statute miles = 93 239 572 square meters

TABLE 1-17 Volume and Capacity Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of volume is the cubic meter.

A. Volume units decimally related to one cubic meter

	Cubic meters (steres) $(m)^3$	Cubic decimeters $(dm)^3$	Cubic centimeters $(cm)^3$	Liters (L)	Centiliters (cL)	Milliliters (mL)	Microliters (μL)
1 cubic meter =	**1**	**1 000**	**1 000 000**	**1 000**	**100 000**	**1 000 000**	$\mathbf{10^{9}}$
1 cubic decimeter =	**0.001**	**1**	**1 000**	**1**	**100**	**1 000**	**1 000 000**
1 cubic centimeter =	**0.000 001**	**0.001**	**1**	**0.001**	**0.1**	**1**	**1 000**
1 liter =	**0.001**	**1**	**1 000**	**1**	**100**	**1 000**	**1 000 000**
1 centiliter =	**0.000 01**	**0.01**	**10**	**0.01**	**1**	**10**	**10 000**
1 milliliter =	**0.000 001**	**0.001**	**1**	**0.001**	**0.1**	**1**	**1 000**
1 microliter =	$\mathbf{10^{-9}}$	**0.000 001**	**0.001**	**0.000 001**	**0.000 1**	**0.001**	**1**

B. Nonmetric volume units (with cubic meter and liter equivalents)

	Cubic meters (steres) $(m)^3$	Liters (L)	Cubic inches $(in)^3$	Cubic feet $(ft)^3$	Cubic yards $(yd)^3$	Barrels (U.S.A.) (bbl)	Acre-Feet (acre-ft)	Cubic miles $(mi)^3$
1 cubic meter =	**1**	**1 000**	$6.102\ 374\ 41 \times 10^{4}$	35.314 666	1.307 950 62	6.289 810 97	$8.107\ 131\ 94 \times 10^{-4}$	$2.399\ 127\ 59 \times 10^{-10}$
1 liter =	**0.001**	**1**	61.023 744 1	0.035 314 66	$1.307\ 950\ 62 \times 10^{-3}$	$6.289\ 810\ 97 \times 10^{-3}$	$8.107\ 131\ 93 \times 10^{-7}$	$2.399\ 127\ 59 \times 10^{-13}$
1 cubic inch =	$\mathbf{1.638\ 706\ 4 \times 10^{-5}}$	$\mathbf{1.638\ 706\ 4 \times 10^{-2}}$	**1**	**1/1 728** = $5.787\ 037\ 03 \times 10^{-4}$	**1/46 656** = $2.143\ 347\ 05 \times 10^{-5}$	$1.030\ 715\ 32 \times 10^{-4}$	$1.328\ 520\ 90 \times 10^{-8}$	$3.931\ 465\ 73 \times 10^{-15}$
1 cubic foot =	$2.831\ 684\ 66 \times 10^{-2}$	**28.316 846 592**	**1 728**	**1**	**1/27** = 0.037 037	0.178 107 61	**1/43 560** = $2.295\ 684\ 11 \times 10^{-5}$	$6.793\ 572\ 78 \times 10^{-12}$
1 cubic yard =	0.764 554 86	764.55 485 8	**46 656**	**27**	**1**	4.808 905 38	$6.198\ 347\ 11 \times 10^{-4}$	$1.834\ 264\ 65 \times 10^{-10}$
1 barrel (U.S.A) =	0.158 987 29	158.987 294	**9 702**	5.614 583 33	0.207 947 53	**1**	$1.288\ 930\ 98 \times 10^{-4}$	$3.814\ 308\ 05 \times 10^{-11}$
1 acre-foot =	1.233 481 84	$1.233\ 481\ 84 \times 10^{6}$	$7.527\ 168\ 00 \times 10^{7}$	**43 560**	1 613 333 33	7 758.367 34	**1**	$2.959\ 280\ 30 \times 10^{-7}$
1 cubic mile =	$4.168\ 181\ 83 \times 10^{9}$	$4.168\ 181\ 83 \times 10^{12}$	$2.543\ 580\ 61 \times 10^{14}$	$\mathbf{1.471\ 979\ 52 \times 10^{11}}$	$\mathbf{5.451\ 776 \times 10^{9}}$	$26.217\ 074\ 9 \times 10^{9}$	**3 379 200**	**1**

TABLE 1-17 Volume and Capacity Conversion Factors (*Continued*)

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of volume is the cubic meter.

C. United States liquid capacity measures (with liter equivalents)

	Liters (L)	Gallons (U.S. gal)	Quarts (U.S. qt)	Pints (U.S. pt)	Gills (U.S. gi)	Fluid ounces (U.S. floz)	Fluidrams (U.S. fldr)	Minims (U.S. minim)
1 liter =	**1**	0.264 172 05	1.056 688	2.113 376	8.453 506	33.814 023	270.512 18	16 230.73
1 gallon, U.S. =	3.785 411 8	**1**	**4**	**8**	**32**	**128**	**1 024**	**61 440**
1 quart, U.S. =	**0.946 352 946**	**1/4 = 0.25**	**1**	**2**	**8**	**32**	**256**	**15 360**
1 pint, U.S. =	0.473 176 5	**1/8 = 0.125**	**1/2 = 0.5**	**1**	**4**	**16**	**128**	**7 680**
1 gill, U.S. =	0.118 294 1	**1/32 = 0.031 25**	**1/8 = 0.125**	**1/4 = 0.25**	**1**	**4**	**32**	**1 920**
1 fluid ounce, U.S. =	2.957 353 × 10^{-2}	**1/128 = 0.007 812 5**	**1/32 = 0.031 25**	1/16 = 0.062 5	**1/4 = 0.25**	**1**	**8**	**480**
1 fluidram, U.S. =	3.696 691 2 × 10^{-3}	**1/102 4 = 9.765 625 × 10^{-4}**	**1/256 = 3.906 25 × 10^{-3}**	**1/128 = 0.007 812 5**	**1/32 = 0.031 25**	**1/8 = 0.125**	**1**	**60**
1 minim, U.S. =	6.161 152 × 10^{-5}	**1/61 440 =** 1.627 604 $\underline{16}$ × 10^{-5}	**1/15 360 =** 6.510 416 $\underline{66}$ × 10^{-5}	**1/7 680 =** 1.302 083 3$\underline{3}$ × 10^{-4}	**1/1 920 =** 5.208 333 $\underline{3}$ × 10^{-4}	**1/480 =** 2.083 333 $\underline{3}$ × 10^{-3}	**1/60 =** 0.016 666 $\underline{6}$	**1**

D. British Imperial liquid capacity measures (with liter equivalents)

	Liters (L)	Gallons (U.K. gal)	Quarts (U.K. qt)	Pints (U.K. pt)	Gills (U.K. gi)	Fluid ounces (U.K. floz)	Fluidrams (U.K. fldr)	Minims (U.K. minim)
1 liter =	**1**	0.219 969 2	0.879 876 6	1.759 753	7.039 018	35.195 06	281.560 5	16 893.63
1 gallon, U.K. =	4.546 092	**1**	**4**	**8**	**32**	**160**	**1 280**	**76 800**
1 quart, U.K. =	1.136 523	**1/4 = 0.25**	**1**	**2**	**8**	**40**	**320**	**19 200**
1 pint, U.K. =	0.568 261 5	**1/8 = 0.125**	**1/2 = 0.5**	**1**	**4**	**20**	**160**	**9 600**
1 gill, U.K. =	0.142 065 4	**1/32 = 0.031 25**	**1/8 = 0.125**	**1/4 = 0.25**	**1**	**5**	**40**	**2 400**
1 fluid ounce, U.K. =	2.841 307 × 10^{-2}	**1/160 = 0.006 25**	**1/40 = 0.025**	**1/20 = 0.05**	**1/5 = 0.2**	**1**	**8**	**480**
1 fluidram, U.K. =	3.551 634 × 10^{-3}	**1/1280 = 7.812 5 × 10^{-4}**	**1/320 = 0.003 125**	**1/160 = 0.006 25**	**1/40 = 0.025**	**1/8 = 0.125**	**1**	**60**
1 minim, U.K. =	5.919 391 × 10^{-5}	**1/76 800 =** 1.302 083 3$\underline{3}$ × 10^{-5}	**1/19 200 =** 5.208 333 3$\underline{3}$ × 10^{-5}	**1/9 600 =** 1.041 666 6$\underline{6}$ × 10^{-4}	**1/2 400 =** 4.166 666 6$\underline{6}$ × 10^{-4}	**1/480 =** 2.083 333 3$\underline{3}$ × 10^{-3}	**1/60 =** 0.016 666 $\underline{66}$	**1**

E. United States and British dry capacity measures (with liter equivalents)

		U.S. dry measures				British dry measures			
	Liters (L)	Bushels (U.S. bu)	Pecks (U.S. peck)	Quarts (U.S. qt)	Pints (U.S. pt)	Bushels (U.K. bu)	Pecks (U.K. peck)	Quarts (U.K. qt)	Pints (U.K. pt)
1 liter =	**1**	0.028 377 59	0.113 510 37	0.908 082 99	1.816 165 98	0.027 496 1	0.109 984 6	0.879 876 6	1.759 753 4
1 bushel, U.S. =	35.239 070	**1**	**4**	**32**	**64**	0.968 938 7	3.875 754 9	31.006 04	62.012 08
1 peck, U.S. =	8.809 767 5	**1/4 = 0.25**	**1**	**8**	**16**	0.242 234 7	0.968 938 7	7.751 509	15.503 02
1 quart, U.S. =	1.101 220 9	**1/32 = 0.031 25**	**1/8 = 0.125**	**1**	**2**	0.030 279 34	0.121 117 3	0.968 938 7	1.937 878
1 pint, U.S. =	0.550 610 5	**1/64 = 0.015 625**	**1/16 = 0.062 5**	**1/2 = 0.5**	**1**	0.015 139 67	0.060 558 67	0.484 469 3	0.968 938 7
1 bushel, U.K. =	36.368 73	1.032 057	4.128 228	33.025 82	66.051 65	**1**	**4**	**32**	**64**
1 peck, U.K. =	9.092 182	0.258 014 3	1.032 057	8.256 456	16.512 91	**1/4 = 0.25**	**1**	**8**	**16**
1 quart, U.K. =	1.136 523	0.032 251 78	0.129 007 1	1.032 057	2.064 114.2	**1/32 = 0.031 25**	**1/8 = 0.125**	**1**	**2**
1 pint, U.K. =	0.568 261 4	0.016 125 89	0.064 503 6	0.516 028 4	1.032 057	**1/64 = 0.015 625**	**1/64 = 0.062 5**	**1/2 = 0.5**	**1**

Exact conversion: 1 dry pint, U.S. = **33.600 312 5** enblc inches

F. Other volume and capacity units

1 barrel, U.S. (used for petroleum, etc.) = **42** gallons = 0.158.987 296 cubic meter
1 barrel ("old barrel") = **31.5** gallons = 0.119 240 cubic meter
1 board foot = **144** cubic inches = 2.359 737 × 10^{-3} cubic meter
1 cord = **128** cubic feet = 3.624 556 cubic meters
1 cord foot = **16** cubic feet = 0.453 069.5 cubic meter
1 cup = **8** fluid ounces, U.S. = 2.365 882 × 10^{-4} cubic meter
1 gallon (Canadian, liquid) = 4.546 090 × 10^{-3} cubic meter
1 perch (volume) = 24.75 cubic feet = 0.700 842 cubic meter
1 stere = **1** cubic meter
1 tablespoon = **0.5** fluid ounce, U.S. = 1.478 677 × 10^{-5} cubic meter
1 teaspoon = **1/6** fluid ounce, U.S. = 4.928 922 × 10^{-6} cubic meter
1 ton (register ton) = **100** cubic feet = 2.831 684 66 cubic meters

TABLE 1-18 Mass Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of mass is the kilogram.

A. Mass units decimally related to one kilogram

	Kilograms (kg)	Tonnes (metric tons)	Grams (g)	Decigrams (dg)	Centigrams (cg)	Milligrams (mg)	Micrograms (μg)
1 kilogram =	**1**	**0.001**	**1 000**	**10 000**	**100 000**	**1 000 000**	**10^9**
1 tonne =	**1 000**	**1**	**1 000 000**	**10^7**	**10^8**	**10^9**	**10^{12}**
1 gram =	**0.001**	**0.000 001**	**1**	**10**	**100**	**1 000**	**1 000 000**
1 decigram =	**0.000 1**	**10^{-7}**	**0.1**	**1**	**10**	**100**	**100 000**
1 centigram =	**0.000 01**	**10^{-8}**	**0.01**	**0.1**	**1**	**10**	**10 000**
1 milligram =	**0.000 001**	**10^{-9}**	**0.001**	**0.01**	**0.1**	**1**	**1 000**
1 microgram =	**10^{-9}**	**10^{-12}**	**0.000 001**	**0.000 01**	**0.000 1**	**0.001**	**1**

B. Nonmetric mass units less than one pound-mass (with gram equivalents)

	Grams (g)	Avoirdupois ounces-mass (oz_m, avdp)	Troy ounces-mass (oz_m, troy)	Avoirdupois drams (dr avdp)	Apothecary drams (dr apoth)	Pennyweights (dwt)	Grains (grain)	Scruples (scruple)
1 gram =	**1**	0.035 273 962	0.032 150 747	0.564 383 39	0.257 205 97	0.643 014 93	15.432 358 4	0.771 617 92
1 avdp ounce-mass =	28.349 523 1	**1**	0.911 458 33	**16**	7.291 666 66	18.227 166 7	**437.5**	**21.875**
1 troy ounce-mass =	31.103 476 8	1.097 142 86	**1**	17.554 285 7	**8**	**20**	**480**	**24**
1 avdp dram =	1.771 845 20	**1/16 = 0.062 5**	0.056 966 15	**1**	0.455 729 17	1.139 322 92	**27.343 75**	**1.367 187 5**
1 apothecary dram =	3.887 934 58	0.137 142 857	**1/8 = 0.125**	2.194 285 70	**1**	**2.5**	**60**	**3**
1 pennyweight =	1.555 173 83	0.054 863 162	**1/20 = 0.05**	0.877 714 28	**1/2.5 = 0.4**	**1**	**24**	**1.2**
1 grain =	**0.064 798 91**	**1/437.5** = 2.285 714 29 × 10^{-3}	**1/480** = 0.002 0833 33	3.657 142 85 × 10^{-2}	**1/60** = 0.016 666 66	**1/24** = 0.041 666 66	**1**	**0.05**
1 scrople =	1.295 078 20	4.571 428 58 × 10^{-2}	**1/24** = 0.041 666 66	0.731 428 57	**1/3** = 0.333 333 33	**5/6** = 0.833 333 33	**20**	**1**

C. Nonmetric mass units of one pound-mass and greater (with kilogram equivalents)

	Kilograms (kg)	Long tons (long ton)	Short tons (short ton)	Long hundredweights (long cwt)	Short hundredweights (short cwt)	Slugs (slug)	Avoirdupois pounds-mass (lb_m, avdp)	Troy pounds-mass (lb_m, troy)
1 kilogram =	**1**	9.842 065 28 × 10^{-1}	1.102 311 31 × 10^{-3}	1.968 411 31 × 10^{-2}	2.204 622 62 × 10^{-2}	0.068 521 77	2.204 622 62	2.679 228 89
1 long ton =	1 016.046 9	**1**	**1.12**	**20**	**22.4**	69.621 329	**2 240**	2 722 222 22
1 short ton =	**907.184 74**	**200/224** = 0.892 857 14	**1**	**4 000/224** = 17.857 142 9	**20**	62.161 901	**2 000**	2 430.555 55
1 long hundredweight =	50.802 345 4	**0.05**	**0.056**	**1**	**1.12**	3.481 066 4	**112**	136.111 11$\underline{1}$
1 short hundredweight =	**45.359 237**	**10/224** = 0.044 642 86	**0.05**	**100/112** = 0.892 857 14	**1**	3.108 095 0	**100**	121.527 7$\underline{77}$
1 slug =	14.593 903	0.014 363 41	0.016 087 02	0.287 268 3	0.321 740 5	**1**	32.174 05	39.100 406
1 avdp pound-mass =	**0.453 592 37**	**1/2 240** = 4.464 285 71 × 10^{-1}	**0.000 5**	**1/1 12** = 8.928 571 43 × 10^{-3}	**0.01**	3.108 095 0 × 10^{-2}	**1**	1.215 277 7$\underline{77}$
1 troy pound-mass =	0.373 241 72	3.673 469 37 × 10^{-1}	4.114 285 70 × 10^{-1}	7.346 938 79 × 10^{-3}	8.228 571 45 × 10^{-3}	0.025 575 18	0.822 857 14	**1**

Exact conversions: 1 long ton = **1 016.046 908 8** kilograms
1 troy pound-mass = **0.373 241 721 6** kilogram

D. Other mass units

1 assay ton = 29.166 667 grams
1 carat (metric) = **200** milligrams
1 carat (troy weight) = **3 1/6** grains = 205.196 55 milligrams
1 myriagram = **10** kilograms
1 quintal = **100** kilograms
1 stone = **14** pounds. avdp = **6.350 293 18** kilograms

TABLE 1-19 Time Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of time is the second.

A. Time units of one second and less

	Seconds (s)	Milliseconds (ms)	Microseconds (μs)		Picoseconds (ps)
1 second =	**1**	**1 000**	**1 000 000**	$\mathbf{10^{9}}$	$\mathbf{10^{12}}$
1 millisecond =	**0.001**	**1**	**1 000**	**1 000 000**	$\mathbf{10^{9}}$
1 microsecond =	**0.000 001**	**0.001**	**1**	**1 000**	**1 000 000**
1 nanosecond =	$\mathbf{10^{-9}}$	**0.000 001**	**0.001**	**1**	**1 000**
1 picosecond =	$\mathbf{10^{-12}}$	$\mathbf{10^{-9}}$	**0.000 001**	**0.001**	**1**

B. Time units of one second and greater

	Mean solar seconds (s)	Mean solar minutes (min)	Mean solar hours (h)	Mean solar days (d)	Mean solar weeks (w)	Calendar (Gregorian) year (yr)
1 second =	**1**	**1/60** = 0.016 666 $\underline{6}$	**1/3 600** = 0.000 277 $\underline{7}$	**1/86 400** = 1.157 407 $\underline{407} \times 10^{-5}$	**1/604 800** = $1.653\ 439\ 15 \times 10^{-6}$	$3.168\ 873\ 85 \times 10^{-8}$
1 minute =	**60**	**1**	**1/60** = 0.016 666 $\underline{6}$	**1/1 440** = 0.000 694 4$\underline{4}$	**1/10 080** = $9.920\ 634\ 92 \times 10^{-5}$	$1.901\ 324\ 31 \times 10^{-6}$
1 hour =	**3 600**	**60**	**1**	**1/24** = 0.041 666 $\underline{6}$	**1/168** = $5.952\ 380\ 95 \times 10^{-3}$	$1.140\ 794\ 50 \times 10^{-4}$
1 day =	**86 400**	**1 440**	**24**	**1**	**1/7 = 0.142 857 14**	$2.737\ 907\ 00 \times 10^{-3}$
1 week =	**604 800**	**10 080**	**168**	**7**	**1**	$1.916\ 534\ 90 \times 10^{-2}$
1 calendar year = (Gregorian)	31 556 952	525 949.2	8 765.82	365.242 5	52.117 5	**1**

NOTES: The conventional calendar year of 365 days can be used in rough calculations only; the modern calendar is based on the Gregorian year of 365.2425 mean solar days, the value chosen by Pope Gregory XIII in 1582. This value requires that a leap-year day be introduced every four years as February 29, except that centennial years (1900, 2000, etc) are leap years only when divisible by 400. The remaining difference between the Gregorian year and the tropical year (see below) introduces an error of 1 day in 3300 years.

The tropical year is the interval between successive vernal equinoxes and has been defined by the International Astronomical Union for noon of January 1, 1900 as 31 556 925.974 7 seconds = 365.242 198 79 mean solar days. The tropical year decreases by approximately 5.3 milliseconds per year.

The sidereal year is the interval between successive returns of the sun to the direction of the same star. Sidereal time units, given in Table 1-18C, are used primarily in astronomy. The SI second, defined by the atomic process of the cesium atom, is equal to the mean solar second within the limits of their definition.

C. Other time units

1 decade = **10** Gregorian years
1 fortnight = **14** days = **1 209 600** seconds
1 century = **100** Gregorian years
1 millennium = **1000** Gregorian years
1 sidereal year = 366.256 4 sidereal days = 31 558 149.8 seconds
1 sidereal day = 86 164.091 seconds
1 sidereal hour = 3 590.170 seconds
1 sidereal minute = 59.836 17 seconds
1 sidereal second = 0.997 269 6 second
1 shake = $\mathbf{10^{-8}}$ seconds

TABLE 1-20 Velocity Conversion Factors

The SI unit of velocity is the meter per second.

	Meters per second (m/s)	Kilometers per hour (km/h)	Statute miles per hour (mi/h)	Knots (kn)	Feet per minute (ft/min)	Feet per second (ft/s)	Inches per second (in/s)
1 meter per second =	**1**	**3.6**	2.236 936 29	1.943 844 49	196.850 394	3.280 839 89	39.370 0787
1 kilometer per hour =	**1/3.6** = 0.277 777	**1**	0.621 371 19	0.539 956 80	54.680 664 9	0.911 344 42	10.936 133 0
1 statute mile per hour =	**0.447 04**	**1.609 344**	**1**	0.868 976 24	**88**	**88/60** = 1.466 666	**88/5** = **17.6**
1 knot =	0.514 444	**1.852**	1.150 779 45	**1**	101.268 592	1.687 780 99	20.253 718 4
1 foot per minute =	**0.005 08**	**0.018 288**	0.011 363	$9.874\ 730\ 01 \times 10^{-3}$	**1**	**1/60** = 0.016 666	**1/5** = **0.2**
1 foot per second =	**0.304 8**	**1.097 28**	0.681 818	0.592 483 80	**60**	**1**	**12**
1 inch per second =	**0.025 4**	**0.091 44**	0.056 818	0.049 373 65	**5**	**1/12** = 0.083 333	**1**

NOTE: The velocity of light in vacuum, c = 299 792 458 meters per second = 670 616 629 statute miles per hour
= 186 282.397 statute miles per second
= 0.983 571 056 feet per nanosecond

Other velocity units

1 foot per hour = $8.466\ 667 \times 10^{-5}$ meter per second
1 statute mile per minute = **26.822 4** meters per second
1 statute mile per second = **1 609.344** meters per second

TABLE 1-21 Density Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of density is the kilogram per cubic meter.

A. Density units decimally related to one kilogram per cubic meter

	Kilograms per cubic meter (kg/m^3)	Tonnes per cubic meter (t/m^3)	Grams per cubic meter (g/m^3)	Grams per liter (g/L)	Milligrams per liter (mg/L)	Micrograms per milliliter (µg/mL)
1 kilogram per cubic meter =	**1**	**0.001**	**1 000**	**1**	**1 000**	**1 000**
1 tonne per cubic meter =	**1 000**	**1**	**1 000 000**	**1 000**	**1 000 000**	**1 000 000**
1 gram per cubic meter =	**0.001**	**0.000 001**	**1**	**0.001**	**1**	**1**
1 gram per liter =	**1**	**0.001**	**1 000**	**1**	**1 000**	**1 000**
1 milligram per liter =	**0.001**	**0.000 001**	**1**	**0.001**	**1**	**1**
1 microgram per milliliter =	**0.001**	**0.000 001**	**1**	**0.001**	**1**	**1**

B. Nonmetric density units (with kilogram per cubic meter equivalents)

	Kilograms per cubic meter (kg/m^3)	Short tons per cubic mile (short tons/mi^3)	Avoirdupois pounds per acrefoot (lb avdp/acre-ft)	Avoirdupois pounds per cubic foot (lb avdp/ft^3)	Avoirdupois pounds per cubic inch (lb avdp/in^3)	Avoirdupois ounces per U.S. quart (oz advp/U.S. qt)	Avoirdupois drams per U.S. fluid ounce (dr advp/U.S. floz)	Grains per U.S. fluid ounce (grain/U.S. floz)
1 kilogram per cubic meter =	**1**	4 594 934	2 719.362 0	$6.242\ 796\ 1 \times 10^{-2}$	$3.612\ 729\ 20 \times 10^{-5}$	$3.338\ 161\ 6 \times 10^{-2}$	$1.669\ 080\ 82 \times 10^{-2}$	0.456 389 28
1 short ton per cubic mile =	$2.176\ 451\ 9 \times 10^{-7}$	**1**	$5.918\ 560\ 5 \times 10^{-4}$	$1.358\ 7145 \times 10^{-8}$	$7.862\ 931\ 3 \times 10^{-12}$	$7.265\ 348\ 2 \times 10^{-9}$	$3.632\ 674\ 1 \times 10^{-9}$	$9.933\ 0931\ 1 \times 10^{-8}$
1 avdp pound per acrefoot =	$3.677\ 333\ 2 \times 10^{-4}$	1 689.600 0	**1**	$2.295\ 684\ 1 \times 10^{-5}$	$1.328\ 520\ 9 \times 10^{-8}$	$1.227\ 553\ 2 \times 10^{-5}$	$6.137\ 766\ 2 \times 10^{-6}$	$1.678\ 295\ 5 \times 10^{-4}$
1 avdp pound per cubic foot =	16.018 463 4	73 598 976	**43 560**	**1**	**1/1 728** = $5.787\ 03\underline{7}\ \underline{03} \times 10^{-4}$	0.534 722 $\underline{2}$	0.267 361 $\underline{1}$	7.310 655 0
1 avdp pound per cubic inch =	27 679.905	$1.271\ 790\ 4 \times 10^{11}$	**75 271 680**	**1 728**	**1**	**924**	**462**	12 632.812
1 avdp ounce per U.S. quart =	29.956 608	$1.376\ 395\ 5 \times 10^{8}$	81 462.86	1.870 130 0	$1.082\ 251\ 1 \times 10^{-3}$	**1**	**0.5**	13.671 874
1 avdp dram per U.S. fluid ounce =	59.913 216	$2.752\ 793\ 0 \times 10^{8}$	162 925.72	3.740 259 8	$2.164\ 502\ 3 \times 10^{-3}$	**2**	**1**	27.343 748
1 grain per U.S. fluid ounce =	2.191 111 9	10 067 357	5 958.426 3	0.136 786 65	$7.915\ 894\ 0 \times 10^{-5}$	0.073 142 86	0.036 571 43	**1**

C. Other density units

1 grain per gallon, U.S. = 17.118 06 grams per cubic meter
1 gram per cubic centimeter = **1 000** kilograms per cubic meter
1 avdp ounce per gallon, U.S. = 7.489 152 kilograms per cubic meter
1 avdp ounce per cubic inch = 1 729.994 kilograms per cubic meter
1 avdp pound per gallon, U.S. = 119.826 4 kilograms per cubic meter
1 slug per cubic foot = 515.379 kilograms per cubic meter
1 long ton per cubic yard = 1 328.939 kilograms per cubic meter

TABLE 1-22 Force Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of force is the newton (N).

	Newtons (N)	Kips (kip)	Slugs-force ($slug_f$)	**Kilograms-force (kilopond) (kg_f)**	Avoirdupois pounds-force (lb_f avdp)	Avoirdupois ounces-force (oz_f advp)	Poundals (pdl)	Dynes (dyn)
1 newton =	**1**	2.248 089 43 × 10^{-4}	6.987 275 24 × 10^{-3}	0.101 971 62	0.224 808 94	3.596 943 09	7.233 014 2	**100 000**
1 kip =	444 8.221 62	**1**	31.080 949	453.592 370	**1 000**	**16 000**	32 174.05	444 822 162
1 slug-force =	143.117 305	0.032 174 05	**1**	14.593 903	32.174 05	514 784 80	1 035.169 5	14 311 730
1 kilogram force (kilopond) =	**9.806 650**	2.204 622 62 × 10^{-3}	6.852 176 3 × 10^{-2}	**1**	2.204 622 62	35.273 961 9	70 931 638 4	**980 665**
1 avdp pound force =	4.448 221 62	**0.001**	3.108 094 88 × 10^{-2}	0.453 592 37	**1**	**16**	32.174 05	444 822.162
1 avdp ounce force =	0.278 013 85	**1/16 000 = 0.000 062 5**	1.942 559 30 × 10^{-3}	2.834 952 3 × 10^{-2}	**1/16 = 0.062 5**	**1**	2.010 878 03	27 801.385
1 poundal =	0.138 254 95	3.108 094 9 × 10^{-5}	9.660 253 9 × 10^{-4}	0.140 980 81	0.031 080 95	0.497 295 18	**1**	13 825.495
1 dyne =	**0.000 01**	2.248 089 43 × 10^{-8}	6.987 275 24 × 10^{-8}	1.019 716 21 × 10^{-6}	2.248 089 43 × 10^{-6}	3.596 943 10 × 10^{-5}	7.233 014 2 × 10^{-5}	**1**

The exact conversion is 1 avdp pound-force = **4.448 221 615 260 5** newtons.

TABLE 1-23 Pressure/Stress Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of pressure or stress is the pascal (Pa).

A. Pressure units decimally related to one pascal

	Pascals (Pa)	Bars (bar)	Decibars (dbar)	Millibars (mbar)	Dynes per square centimeter (dyn/cm^2)
1 pascal =	**1**	**0.000 01**	**0.000 1**	**0.01**	**10**
1 bar =	**100 000**	**1**	**10**	**1 000**	**1 000 000**
1 decibar =	**10 000**	**0.1**	**1**	**100**	**100 000**
1 millibar =	**100**	**0.001**	**0.01**	**1**	**1 000**
1 dyne per square centimeter =	**0.1**	**0.000 001**	**0.000 01**	**0.001**	**1**

B. Pressure units decimally related to one kilogram-force per square meter (with pascal equivalents)

	Kilograms-force per square meter (kg_f/m^2)	Kilograms-force per square centimeter (kg_f/cm^2)	Kilograms-force per square millimeter (kg_f/mm^2)	Grams-force per square centimeter (g_f/cm^2)	Pascals (Pa)
1 kilogram-force per square meter =	**1**	**0.000 1**	**0.000 001**	**0.1**	**9.806 65**
1 kilogram-force per square centimeter =	**10 000**	**1**	**0.01**	**1 000**	**98 066.5**
1 kilogram-force per square millimeter =	**1 000 000**	**100**	**1**	**100 000**	**9 806 650**
1 gram-force per square centimeter =	**10**	**0.001**	**0.000 01**	**1**	**98.066 5**
1 pascal =	0.101 971 62	$1.019\,7162 \times 10^{-5}$	$1.019\,716\,2 \times 10^{-7}$	$1.019\,716\,2 \times 10^{-2}$	1

NOTE: **1** atmosphere (technical) = 1 kilogram-force per square centimeter = **98 066.5** pascals.

C. Pressure units expressed as heights of liquid (with pascal equivalents)

	Millimeters of mercury at 0°C (mmHg, 0°C)	Centimeters of mercury at 60°C (cmHg, 60°C)	Inches of mercury at 32°F (inHg, 32°F)	Inches of mercury at 60°F (inHg, 60°F)	Centimeters of water at 4°C (cmH_2O, 4°C)	Inches of water at 60°F (inH_2O, 60°F)	Feet of water at 39.2°F (ftH_2O, 39.2°F)	Pascals (Pa)
1 millimeter of mercury, 0°C =	**1**	0.100 282	0.039 370 1	0.039 481 3	1.359 548	0.535 775 6	0.044 604 6	133.322 4
1 centimeter of mercury, 60°C =	9.971 830	**1**	0.392 591 9	0.393 700 8	13.557 18	5.342 664	0.444 789 5	1 329.468
1 inch of mercury, 32°F =	**25.4**	2.547 175	**1**	1.002 824 8	34.532 52	13.608 70	1.132 957	3 386.389
1 inch of mercury, 60°C =	25.328 45	**2.54**	0.997 183 1	**1**	34.435 25	13.570 37	1.129 765	3 376.85
1 centimeter of water, 4°C =	0.735 539	0.073 762	0.028 958	0.029 040 0	**1**	0.394 083 8	0.032 808 4	98.063 8
1 inch of water, 60°F =	1.866 453	0.187 173	0.073 482	0.073 690 0	2.537 531	**1**	0.083 252 4	248.840
1 foot of water, 39.2°F =	22.419 2	2.248 254	0.882 646	0.885 139	30.479 98	12.011 67	**1**	2 988.98
1 pascal =	$7.500\ 615 \times 10^{-3}$	$7.521\ 806 \times 10^{-4}$	$2.952\ 998 \times 10^{-4}$	$2.961\ 34 \times 10^{-4}$	$1.019\ 74 \times 10^{-2}$	$4.018\ 65 \times 10^{-3}$	$3.345\ 62 \times 10^{-4}$	**1**

NOTE: 1 torr = 1 millimeter of mercury at 0°C = 133.322 4 pascals.

D. Nonmetric pressure units (with pascal equivalents)

	Atmospheres (atm)	Avoirdupois pounds-force per square inch (lb/in^2)	Avoirdupois pounds-force per square foot (lb_f/ft^2, avdp)	Poundals per square foot (pdl/ft^2)	Pascals (Pa)
1 atmosphere =	**1**	14.695 95	2 116.217	68 087.24	**101 325**
1 avdp pound-force per square inch =	$6.804\ 60 \times 10^{-2}$	**1**	**144**	4 633.063	6 894.757
1 avdp pound-force per square foot =	$4.725\ 414 \times 10^{-4}$	**1/144** = 0.006 94$\underline{4}$	**1**	32.174 05	47.880 26
1 poundal per square foot =	$1.468\ 704 \times 10^{-5}$	$2.158\ 399 \times 10^{-4}$	0.031 080 9	**1**	1.488 164
1 pascal =	$9.869\ 233 \times 10^{-6}$	$1.450\ 377 \times 10^{-4}$	0.020 885 4	0.671 968 9	**1**

NOTE: 1 normal atmosphere = 760 torr = **101 325** pascals.

TABLE 1-24 Torque/Bending Moment Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are <u>underlined</u>.) The SI unit of torque is the newton-meter (N · m).

	Newton-meters (N · m)	Kilogram-force-meters (kg_f · m)	Avoirdupois pound-force-feet (lb_f · ft, avdp)	Avoirdupois pound-force-inches (lb_f · in, avdp)	Avoirdupois ounce-force-inches (oz_f · in, avdp)	Dyne-centimeters (dyne · cm)
1 newton-meter =	**1**	0.101 971 6	0.737 562 1	8.850 748 1	141.611 9	**10 000 000**
1 kilogram-force-meter =	**9.806 65**	**1**	7.233 013	86.796 16	1 388.739	**98 066 500**
1 avdp pound-force-foot =	1.355 818	0.138 255 0	**1**	**12**	**192**	13 558 180
1 avdp pound-force-inch =	0.112 984 8	$1.152\ 124 \times 10^{-2}$	**1/12** = 0.083 33<u>3</u>	**1**	**16**	1 129 848
1 avdp ounce-force-inch =	$7.061\ 552 \times 10^{-3}$	$7.200\ 779 \times 10^{-4}$	**1/192** = 0.005 208 <u>3</u>	**1/16 = 0.062 5**	**1**	70 615.52
1 dyne-centimeter =	$\mathbf{10^{-7}}$	$1.017\ 716 \times 10^{-8}$	$7.375\ 621 \times 10^{-8}$	$8.850\ 748 \times 10^{-7}$	$1.416\ 119 \times 10^{-5}$	**1**

TABLE 1-25 Energy/Work Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of energy and work is the joule (J).

A. Energy/work units decimally related to one joule

	Joules (J)	Megajoules (MJ)	Kilojoules (kJ)	Millijoules (mJ)	Microjoules (μJ)	Ergs (erg)
1 joule =	**1**	**0.000 001**	**0.001**	**1 000**	**1 000 000**	$\mathbf{10^{7}}$
1 megajoule =	**1 000 000**	**1**	**1 000**	$\mathbf{10^{9}}$	$\mathbf{10^{12}}$	$\mathbf{10^{13}}$
1 kilojoule =	**1 000**	**0.001**	**1**	**1 000 000**	$\mathbf{10^{9}}$	$\mathbf{10^{10}}$
1 millijoule =	**0.001**	$\mathbf{10^{-9}}$	$\mathbf{10^{-6}}$	**1**	**1 000**	**10 000**
1 microjoule =	**0.000 001**	$\mathbf{10^{-12}}$	$\mathbf{10^{-9}}$	**0.001**	**1**	**10**
1 erg =	$\mathbf{10^{-7}}$	$\mathbf{10^{-13}}$	$\mathbf{10^{-10}}$	**0.000 1**	**0.1**	**1**

NOTE: I watt-second = 1 joule.

B. Energy/work units less than ten joules (with joule equivalents)

	Joules (J)	Foot-poundals (ft · pdl)	Foot-pounds-force (ft · lb_f)	Calories (International Table) (cal, IT)	Calories (thermochemical) (cal, thermo)	Electronvolts (eV)
1 joule =	**1**	23.730 36	0.737 562 1	0.238 845 9	0.239 005 7	$6.241\ 46 \times 10^{18}$
1 foot-poundal =	$4.214\ 011 \times 10^{-2}$	**1**	$3.108\ 095 \times 10^{-2}$	$1.006\ 499 \times 10^{-2}$	$1.007\ 173 \times 10^{-2}$	$2.630\ 16 \times 10^{17}$
1 foot-pound-force =	1.355 818	32.174 05	**1**	0.323 831 6	0.324 048 3	$8.462\ 28 \times 10^{18}$
1 calorie (Int. Tab.) =	**4.186 8**	99.854 27	3.088 025	**1**	1.000 669	$2.613\ 17 \times 10^{19}$
1 calorie (thermo) =	**4.184**	99.287 83	3.085 960	0.999 331 2	**1**	$2.611\ 43 \times 10^{19}$
1 electronvolt =	$1.602\ 19 \times 10^{-18}$	$3.802\ 05 \times 10^{-18}$	$1.181\ 71 \times 10^{-19}$	$3.826\ 77 \times 10^{-20}$	$3.829\ 33 \times 10^{-20}$	**1**

C. Energy/work units greater than ten joules (with joule equivalents)

	Joules (J)	British thermal units, International Table (Btu, IT)	British thermal units, thermochemical (Btu, thermo)	Kilowatthours (kWh)	Horsepower-hours, electrical (hp · h, elec)	Kilocalories, International Table (kcal, IT)	Kilocalories, thermochemical (kcal, thermo)
1 joule =	**1**	$9.478\ 170 \times 10^{-4}$	$9.484\ 516\ 5 \times 10^{-4}$	$\mathbf{1/(3.6 \times 10^{6})}$ $\mathbf{2.77\underline{7} \times 10^{-7}}$	$3.723\ 562 \times 10^{-7}$	$2.388\ 459 \times 10^{-4}$	$2.390\ 057\ 4 \times 10^{-4}$
1 British thermal unit, Int. Tab. =	1 055.056	**1**	1.000 669	$2.930\ 711\ 1 \times 10^{-4}$	$3.928\ 567 \times 10^{-4}$	0.251 995 8	0.252 164 4
1 British thermal unit (thermo) =	1 054.35	0.999 331	**1**	$2.928\ 745 \times 10^{-4}$	$03.925\ 938 \times 10^{-4}$	0.251 827 2	0.251 995 7
1 kilowatthour =	**3 600 000**	3 412.141	3 414.426	**1**	**1/0.746** = 1.340 482 6	859.845 2	860.420 7
1 horsepower hour, electrical =	**2 685 600**	2 545.457	2 547.162	**0.746**	**1**	641.444 5	641.873 8
1 kilocalorie, Int. Tab. =	**4 186.8**	3.968 320	3.970 977	**0.001 163**	$1.558\ 981 \times 10^{-3}$	**1**	1.000 669
1 kilocalorie, thermochemical =	**4 184**	3.965 666	3.968 322	0.001 162 $\underline{2}$	$1.557\ 938\ 6 \times 10^{-3}$	0.999 331	**1**

The exact conversion is 1 British thermal unit, International Table = **1 055.055 852 62** joules.

TABLE 1-26 Power Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.) The SI unit of power is the watt (W).

A. Power units decimally related to one watt

	Watts (W)	Megawatts (MW)	Kilowatts (kW)	Milliwatts (mW)	Microwatts (μW)	Picowatts (pW)	Ergs per second (ergs/s)
1 watt =	**1**	**0.000 001**	**0.001**	**1 000**	**1 000 000**	**10^{9}**	**10^{7}**
1 megawatt =	**1 000 000**	**1**	**1 000**	**10^{9}**	**10^{12}**	**10^{15}**	**10^{13}**
1 kilowatt =	**1 000**	**0.001**	**1**	**1 000 000**	**10^{9}**	**10^{12}**	**10^{10}**
1 milliwatt =	**0.001**	**10^{-9}**	**0.000 001**	**1**	**1 000**	**1 000 000**	**10 000**
1 microwatt =	**0.000 001**	**10^{-12}**	**10^{-9}**	**0.001**	**1**	**1 000**	**10**
1 picowatt =	**10^{-9}**	**10^{-15}**	**10^{-12}**	**0.000 001**	**0.001**	**1**	**0.01**
1 erg per second =	**10^{-7}**	**10^{-13}**	**10^{-10}**	**0.000 1**	**0.1**	**100**	**1**

NOTE: 1 watt = 1 joule per second (J/s).

B. Nonmetric power units (with watt equivalents)

	British thermal units (International Table) per hour (Btu/hr, IT)	British thermal units (thermochemical) per minute (Btu/min, thermo)	Avoirdupois foot-pounds-force per second ($ft \cdot lb_f/s$ avdp)	Kilocalories per minute (thermochemical) (kcal/min, thermo)	Kilocalories per second (International Table) (kcal/s, IT)	Horsepower (electrical) (hp, elec)	Horsepower (mechanical) (hp, mech)	Watts (W)
1 British thermal unit (Int. Tab.)-per hour =	**1**	0.016 677 8	0.216 158 1	$4.202\ 740\ 5 \times 10^{-3}$	$6.999\ 883\ 1 \times 10^{-5}$	$3.928\ 567\ 0 \times 10^{-4}$	$3.930\ 148\ 0 \times 10^{-4}$	0.293 071 1
1 British thermal unit (thermo) per minute =	59.959 853	**1**	12.960 810	0.251 995 7	$4.197\ 119\ 5 \times 10^{-3}$	0.023 555 6	0.023 565 1	17.572 50
1 foot-pound-force per second =	4.626 242 6	0.077 155 7	**1**	0.019 442 9	$3.238\ 315\ 7 \times 10^{-4}$	$1.817\ 450\ 4 \times 10^{-3}$	**1/550** = $1.818\ 181\ 8 \times 10^{-3}$	1.355 818
1 kilocalorie per minute (thermo) =	237.939 98	3.968 321 7	51.432 665	**1**	0.016 655 5	0.093 476 3	0.093 513 9	69.733 333
1 kilocalorie per second (Int. Tab.) =	14 285.953	238.258 64	3 088.025 1	60.040 153	**1**	5.612 332 4	5.614 591 1	**4 186.800**
1 horsepower (electrical) =	2 545.457 4	42.452 696	550.221 34	10.697 898	0.178 179 0	**1**	1.000 402 4	**746**
1 horsepower (mechanical) =	2 544.433 4	42.435 618	**550**	10.693 593	0.178 107 4	0.999 597 7	**1**	745.699 9
1 watt =	3.412 141 3	0.056 907 1	0.737.562 1	0.014 340 3	$2.388\ 459\ 0 \times 10^{-4}$	**1/746** = $1.340\ 482\ 6 \times 10^{-3}$	$1.341\ 022\ 0 \times 10^{-3}$	**1**

NOTE: The horsepower (mechanical) is defined as a power equal to **550** foot-pounds-force per second.
Other units of horsepower are:
1 horsepower (boiler) = 9 809.50 watts
1 horsepower (metric) = 735.499 watts
1 horsepower (water) = 746.043 watts
1 horsepower (U.K.) = 745.70 watts
1 ton (refrigeration) = 3 516.8 watts

TABLE 1-27 Temperature Conversions

(Conversions in **boldface** type are exact. Continuing decimals are underlined.)

Celsius (°C) °C = 5(°F–32)/9	Fahrenheit (°F) °F = [9(C°)/5] + 32	Absolute (K) K = °C + 273.15
–273.15	**–459.67**	**0**
–200	**–328**	**73.15**
–180	**–292**	**93.15**
–160	**–256**	**113.15**
–140	**–220**	**133.15**
–120	**–184**	**153.15**
–100	**–148**	**173.15**
–80	**–112**	**193.15**
–60	**–76**	**213.15**
–40	**–40**	**233.15**
–20	**–4**	**253.15**
–17.77	0	255.372
0	**32**	**273.15**
5	**41**	**278.15**
10	**50**	**283.15**
15	**59**	**288.15**
20	**68**	**293.15**
25	**77**	**298.15**
30	**86**	**303.15**
35	**95**	**308.15**
40	**104**	**313.15**
45	**113**	**318.15**
50	**122**	**323.15**
55	**131**	**328.15**
60	**140**	**333.15**
65	**149**	**338.15**
70	**158**	**343.15**
75	**167**	**348.15**
80	**176**	**353.15**
85	**185**	**358.15**
90	**194**	**363.15**
95	**203**	**368.15**
100	**212**	**373.15**
105	**221**	**378.15**
110	**230**	**383.15**
115	**239**	**378.15**
120	**248**	**393.15**
140	**284**	**413.15**
160	**320**	**433.15**
180	**356**	**453.15**
200	**392**	**473.15**
250	**482**	**523.15**
300	**572**	**573.15**
350	**662**	**623.15**
400	**752**	**673.15**
450	**842**	**723.15**
500	**932**	**773.15**
1 000	**1 832**	**1 273.15**
5 000	**9 032**	**5 273.15**
10 000	**18 032**	**10 273.15**

NOTE: Temperature in kelvins equals temperature in degrees Rankine divided by 1.8. [K = °R/**1.8**].

TABLE 1-28 Light Conversion Factors

(Exact conversions are shown in **boldface** type. Repeating decimals are underlined.)

A. Luminance units. The SI unit of luminance is the candela per square meter (cd/m²).

	Candelas per square meter (cd/m²)	Candelas per square foot (cd/ft²)	Candelas per square inch (cd/in²)	Apostilbs (asb)	Stilbs (sb)	Lamberts (L)	Footlamberts (fL)
1 candela per square meter =	**1**	**0.092 903 04**	$\mathbf{6.451\ 6 \times 10^{-4}}$	π = 3.141 592 65	**0.000 1**	**(0.000 1)** π = $3.141\ 592\ 65 \times 10^{-4}$	0.291 863 51
1 candela per square foot =	10.763 910 4	**1**	**1/144** = 0.006 944 44	33.815 821 8	$1.076\ 391\ 04 \times 10^{-3}$	$3.381\ 582\ 18 \times 10^{-3}$	π = 3.141 592 65
1 candela per square inch =	1 550.003 1	**144**	**1**	4 869.478 4	0.155 000 31	0.486 947 84	452.389 342
1 apostilb =	**1/π** = 0.318 309 89	0.029 571 96	$2.053\ 608\ 06 \times 10^{-4}$	**1**	$3.183\ 098\ 86 \times 10^{-5}$	**0.000 1**	**0.092 903 04**
1 stilb =	**10 000**	**929.030 4**	**6.451 6**	31 415.926 5	**1**	π = 3.141 592 65	2 918.635
1 lambert =	**10 000/π** = 3 183.098 86	295.719 561	2.053 608 06	**10 000**	**1/π** = 0.318 309 89	**1**	**929.030 4**
1 footlambert =	3.426 259 1	**1/π** = 0.318 309 89	$2.210\ 485\ 32 \times 10^{-3}$	10.763 910 4	$3.426\ 259\ 1 \times 10^{-4}$	$1.076\ 391\ 03 \times 10^{-3}$	**1**

NOTE: 1 nit (nt) = 1 candela per square meter (cd/m²).
1 stilb (sb) = 1 candela per square centimeter (cd/cm²).

B. Illuminance units. The SI unit of illuminance is the lux (lux).

	Luxes (lx)	Phots (ph)	Footcandles (fc)	Lumens per square inch (lm/in²)
1 lux =	**1**	**0.000 1**	**0.092 903 04**	$\mathbf{6.451\ 6 \times 10^{-4}}$
1 phot =	**10 000**	**1**	**929.030 4**	**6.451 6**
1 footcandle =	10.763 910 4	$1.076\ 391\ 04 \times 10^{-3}$	**1**	**1/144** = 0.006 944 44
1 lumen per square inch =	1 550.003 1	0.155 000 31	**144**	**1**

NOTE: **1 lux (lux) = 1 lumen per square meter (lm/m²).**
1 phot (ph) = 1 lumen per square centimeter (lm/cm²).
1 footcandle (fc) = 1 lumen per square foot (lm/ft²).

This table contains similar statements relating the meter, yard, foot, inch, mil, and microinch to each other, that is, conversion factors between the non-SI units as well as to and from the SI unit are given. In all, these tables contain over 1700 such statements. Exact conversion factors are indicated in **boldface** type.

Tabulation Groups. To produce tables that can be contained on individual pages of the handbook, units of a given quantity have been arranged in separate subtabulations identified by capital letters. Each such subtabulation represents a group of units related to each other decimally, by magnitude or by usage. Each subtabulation contains the SI unit,* so equivalent values can be found between units that are tabulated in separate tables. For example, to obtain equivalence between pounds per cubic foot and tonnes per cubic meter, we read from the fourth line of Table 1-21B:

1 pound per cubic foot is equal to 16.018 463 4 kilograms per cubic meter

From the first line of Table 1-21A, we find:

1 kilogram per cubic meter is equal to 0.001 metric ton per cubic meter

Hence,

1 pound per cubic foot is equal to 16.018 463 4 kilograms per cubic meter
= 0.016 018 463 4 metric ton per cubic meter

Use of Conversion Factors. Conversion factors are multipliers used to convert a quantity expressed in a particular unit *(given unit)* to the same quantity expressed in another unit *(desired unit)*. To perform such conversions, the *given unit* is found at the left-hand edge of the conversion table, and the *desired unit* is found at the top of the same table. Suppose, for example, the quantity 1000 feet is to be converted to meters. The given unit, foot, is found in the left-hand edge of the third line of Table 1-15B. The desired unit, meter, is found at the top of the first column in that table. The conversion factor (**0.304 8**, exactly) is located to the right of the given unit and below the desired unit. The given quantity, 1000 feet, is multiplied by the conversion factor to obtain the equivalent length in meters, that is, 1000 feet is $1000 \times 0.304\ 8 = 304.8$ meters.

The general rule is: Find the given unit at the left side of the table in which it appears and the desired unit at the top of the same table; note the conversion factor to the right of the given unit and below the desired unit. Multiply the quantity expressed in the given unit by the conversion factor to find the quantity expressed in the desired unit.

Listings of conversion factors (see Refs. 1 and 7) are often arranged as follows:

To convert from	*To*	*Multiply by*
(Given unit)	(Desired unit)	(Conversion factor)

The equivalences listed in the accompanying conversion tables can be cast in this form by placing the given unit (at the left of each table) under "To convert from," the desired units (at the top of the table) under "To," and the conversion factor, found to the right and below these units, under "Multiply by."

Use of Two Tables to Find Conversion Factors. When the given and desired units do not appear in the same table, the conversion factor between them is found in two steps. The *given unit* is selected at the left-hand edge of the table in which it appears, and an *intermediate conversion factor,* applicable to the SI unit shown at the top of the same table, is recorded. The *desired unit* is then found at the top of another table in which it appears, and another *intermediate conversion factor*, applicable to the SI unit at the left-hand edge of that table, is recorded. The conversion factor between the given and desired units is the product of these two intermediate conversion factors.

*In Tables 1-17C, 1-17D, 1-17E, and 1-18B, a decimal submultiple of the SI unit (the liter and gram, respectively) is listed because it is most commonly used in conjunction with the other units in the respective tables. The procedure for linking the subtables is unchanged.

TABLE 1-29 U.S. Electrical Units Used Prior to 1969, with SI Equivalents

A. Legal units in the U.S. prior to January 1948	
1 ampere (US-INT)	= 0.999 843 ampere (SI)
1 coulomb (US-INT)	= 0.999 843 coulomb (SI)
1 farad (US-INT)	= 0.999 505 farad (SI)
1 henry (US-INT)	= 1.000 495 henry (SI)
1 joule (US-INT)	= 1.000 182 joule (SI)
1 ohm (US-INT)	= 1.000 495 ohm (SI)
1 volt (US-INT)	= 1.000 338 volt (SI)
1 watt (US-INT)	= 1.000 182 watt (SI)
B. Legal units in the U.S. from January 1948 to January 1969	
1 ampere (US-48)	= 1.000 008 ampere (SI)
1 coulomb (US-48)	= 1.000 008 coulomb (SI)
1 farad (US-48)	= 0.999 505 farad (SI)
1 henry (US-48)	= 1.000 495 henry (SI)
1 joule (US-48)	= 1.000 017 joule (SI)
1 ohm (US-48)	= 1.000 495 ohm (SI)
1 volt (US-48)	= 1.000 008 volt (SI)
1 watt (US-48)	= 1.000 017 watt (SI)

For example, it is required to convert 100 cubic feet to the equivalent quantity in cubic centimeters. The given quantity (cubic feet) is found in the fourth line at the left of Table 1-17B. Its intermediate conversion factor with respect to the SI unit is found below the cubic meters to be 2.831 684 66 $\times 10^{-2}$. The desired quantity (cubic centimeters) is found at the top of the third column in Table 1-17A. Its intermediate conversion factor with respect to the SI unit, found under the cubic centimeters and to the right of the cubic meters, is 1 000 000. The conversion factor between cubic feet and cubic centimeters is the product of these two intermediate conversion factors, that is, 1 cubic foot is equal to 2.831 684 66 $\times 10^{-2} \times$ 1 000 000 = 28 316.846 6 cubic centimeters. The conversion from 100 cubic feet to cubic centimeters then yields 100 $\times$ 28 316.846 6 = 2 831 684.66 cubic centimeters.

Conversion of Electrical Units. Since the electrical units in current use are confined to the International System, conversions to or from non-SI units are fortunately not required in modern practice. Conversions to and from the older cgs units, when required, can be performed using the conversions shown in Table 1-9. Slight differences from the SI units occur in the electrical units legally recognized in the United States prior to 1969. These differences involve amounts smaller than that customarily significant in engineering; they are listed in Table 1-29.

BIBLIOGRAPHY

Standards

ANSI/IEEE Std 268; Metric Practice. New York, Institute of Electrical and Electronics Engineers.

Graphic Symbols for Electrical and Electronics Diagrams, IEEE Std 315 (also published as ANSI Std Y32.2). New York, Institute of Electrical and Electronics Engineers.

IEEE Standard Letter Symbols for Units of Measurement, ANSI/IEEE Std 260. New York, Institute of Electrical and Electronics Engineers.

IEEE Recommended Practice for Units in Published Scientific and Technical Work, IEEE Std 268. New York, Institute of Electrical and Electronics Engineers.

Letter Symbols for Quantities Used in Electrical Science and Electrical Engineering; ANSI Std Y10.5. Also published as IEEE Std 280; New York, Institute of Electrical and Electronics Engineers.

SI Units and Recommendations for the Use of Their Multiples and of Certain Other Units; International Standards ISO-1000 (E). Available in the United States from ANSI. New York, American National Standards Institute. Also identified as IEEE Std 322 and ANSI Z210.1.

Collections of Units and Conversion Factors

Encyclopaedia Britannica (see under "Weights and Measures"). Chicago, Encyclopaedia Britannica, Inc.

McGraw-Hill Encyclopedia of Science and Technology (see entries by name of quantity or unit and vol. 20 under "Scientific Notation". New York, McGraw-Hill.

Mohr, Peter J. and Barry N. Taylor, *CODATA*: 2002; Recommended Values of the Fundamental Physical Constants; *Reviews of Modern Physics*, January 2005, vol. 77, no. 1, pp. 1–107, http://www. physics.nist.gov/constants.

National Institute of Standards and Technology Units of Weight and Measure—International (Metric) and U.S. Customary; *NIST Misc. Publ*. 286. Washington, Government Printing Office.

The Introduction of the IAU System of Astronomical Constants into the Astronomical Ephemeris and into the American Ephemeris and Nautical Almanac (Supplement to the American Ephemeris 1968). Washington, United States Naval Observatory, 1966.

The Use of SI Units (The Metric System in the United Kingdom), PD 5686. London, British Standards Institution. See also British Std 350, Part 2, and PD 6203 Supplement 1.

The World Book Encyclopedia (see under "Weights and Measures"). Chicago, Field Enterprises Educational Corporation.

World Weights and Measures, Handbook for Statisticians, Statistical Papers, Series M, No. 21, Publication Sales No. 66, XVII, 3. New York, United Nations Publishing Service.

Books and Papers

Brownridge, D. R.: *Metric in Minutes*. Belmont, CA, Professional Publications, Inc., 1994.

Cornelius, P., de Groot, W., and Vermeulen, R.: Quantity Equations, Rationalization and Change of Number of Fundamental Quantities (in three parts); *Appl. Sci. Res.*, 1965, vol. B12, pp. 1, 235, 248.

IEEE Standard Dictionary of Electrical and Electronics Terms, ANSI/IEEE Std 100-1988. New York, Institute of Electrical and Electronics Engineers, 1988.

Page, C. H.: Physical Entities and Mathematical Representation; *J. Res. Natl. Bur. Standards*, October–December 1961, vol. 65B, pp. 227–235.

Silsbee, F. B.: Systems of Electrical Units; *J. Res. Natl. Bur. Standards*, April–June 1962, vol. 66C, pp. 137–178.

Young, L.: *Systems of Units in Electricity and Magnetism*. Edinburgh, Oliver & Boyd Ltd., 1969.

SECTION 2
ELECTRIC AND MAGNETIC CIRCUITS*

Paulo F. Ribeiro
Professor of Engineering, Calvin College, Grand Rapids, MI, Scholar Scientist, Center for Advanced Power Systems, Florida State University, Fellow, Institute of Electrical and Electronics Engineers

Yazhou (Joel) Liu, PhD
IEEE Senior Member; Thales Avionics Electrical System

CONTENTS

2.1 ELECTRIC AND MAGNETIC CIRCUITS

Definition of Electric Circuit. An *electric circuit* is a collection of electrical devices and components connected together for the purpose of processing information or energy in electrical form. An electric circuit may be described mathematically by ordinary differential equations, which may be linear or

*The authors thank Nate Haveman for assisting with manuscript preparation.

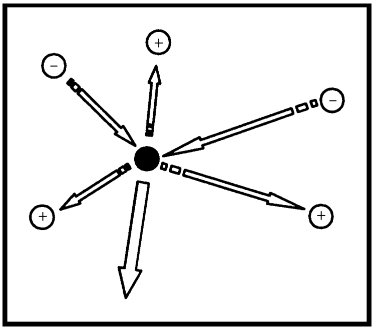


FIGURE 2-1 Electric charges.

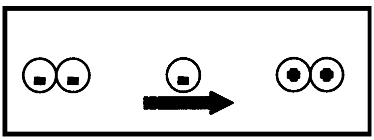


FIGURE 2-2 Electric voltage.

nonlinear, and which may or may not be time varying. The practical effect of this restriction is that the physical dimensions are small compared to the wavelength of electrical signals. Many devices and systems use circuits in their design.

Electric Charge. In circuit theory, we postulate the existence of an indivisible unit of charge. There are two kinds of charge, called *negative* and *positive* charge. The negatively charged particle is called an *electron*. Positive charges may be atoms that have lost electrons, called *ions;* in crystalline structures, electron deficiencies, called *holes,* act as positively charged particles. See Fig. 2-1 for an illustration. In the International System of Units (SI), the unit of charge is the coulomb (C). The charge on one electron is 1.60219×10^{-19} C.

Electric Current. The flow or motion of charged particles is called an *electric current*. In SI units, one of the fundamental units is the ampere (A). The definition is such that a charge flow rate of 1 A is equivalent to 1 C/s. By convention, we speak of current as the flow of positive charges. See Fig. 2-2 for an illustration. When it is necessary to consider the flow of negative charges, we use appropriate modifiers. In an electric circuit, it is necessary to control the path of current flow so that the device operates as intended.

Voltage. The motion of charged particles either requires the expenditure of energy or is accompanied by the release of energy. The voltage, at a point in space, is defined as the work per unit charge (joules/coulomb) required to move a charge from a point of zero voltage to the point in question.

Magnetic and Dielectric Circuits. Magnetic and electric fields may be controlled by suitable arrangements of appropriate materials. Magnetic examples include the magnetic fields of motors, generators, and tape recorders. Dielectric examples include certain types of microphones. The fields themselves are called *fluxes* or *flux fields*. Magnetic fields are developed by magnetomotive forces. Electric fields are developed by voltages (also called *electromotive forces,* a term that is now less common). As with electric circuits, the dimensions for dielectric and magnetic circuits are small compared to a wavelength. In practice, the circuits are frequently nonlinear. It is also desired to confine the magnetic or electric flux to a prescribed path.

2.1.1 Development of Voltage and Current

Sources of Voltage or Electric Potential Difference. A voltage is caused by the separation of opposite electric charges and represents the work per unit charge (joules/coulomb) required to move the charges from one point to the other. This separation may be forced by physical motion, or it may be initiated or complemented by thermal, chemical, magnetic, or radiation causes. A convenient classification of these causes is as follows:

a. Friction between dissimilar substances
b. Contact of dissimilar substances
c. Thermoelectric action
d. Hall effect
e. Electromagnetic induction
f. Photoelectric effect
g. Chemical action

Voltage Effect or Contact Potential. When pieces of various materials are brought into contact, a voltage is developed between them. If the materials are zinc and copper, zinc becomes charged positively and copper negatively. According to the electron theory, different substances possess different tendencies to give up their negatively charged particles. Zinc gives them up easily, and thus, a number of negatively charged particles pass from it to copper. Measurable voltages are observed even between two pieces of the same substance having different structures, for example, between pieces of cast copper and electrolytic copper.

Thomson Effect. A temperature gradient in a metallic conductor is accompanied by a small voltage gradient whose magnitude and direction depend on the particular metal. When an electric current flows, there is an evolution or absorption of heat due to the presence of the thermoelectric gradient, with the net result that the heat evolved in a volume interval bounded by different temperatures is slightly greater or less than that accounted for by the resistance of the conductor. In copper, the evolution of heat is greater when the current flows from hot to cold parts, and less when the current flows from cold to hot. In iron, the effect is the reverse. Discovery of this phenomenon in 1854 is credited to Sir William Thomson (Lord Kelvin), an English physicist.

The Thomson effect is defined by

$$q = \rho J^2 - \mu J \frac{dT}{dx}$$

where q is the heat production per unit volume, ρ is the resistivity of the material, J is the current density, μ is the Thomson coefficient, and dx/dT is the temperature gradient.

Peltier Effect. When a current is passed across the junction between two different metals, an evolution or an absorption of heat takes place. This effect is different from the evolution of heat described by ohmic (i^2r) losses. This effect is reversible, heat being evolved when current passes one way across the junction, and absorbed when the current passes in the opposite direction. The junction is the source of a Peltier voltage. When current is forced across the junction against the direction of the voltage, a heating action occurs. If the current is forced in the direction of the Peltier voltage, the junction is cooled. Refrigerators are constructed using this principle. Since the Joule effect (see Sec. 2.1.8) produces heat in the conductors leading to the junction, the Peltier cooling must be greater than the Joule effect in that region for refrigeration to be successful. This phenomenon was discovered by Jean Peltier, a French physicist, in 1834.

The Peltier effect is defined by

$$Q = \prod_{AB} . I$$

where Q is the heat absorption per unit time, $Q = \prod_{AB}$ is the Peltier coefficient, and I is the current.

Seebeck Effect. When a closed electric circuit is made from two different metals, two (or more) junctions will be present. If these junctions are maintained at different temperatures, within certain ranges, an electric current flows. If the metals are iron and copper, and if one junction is kept in ice while the other is kept in boiling water, current passes from copper to iron across the hot junction. The resulting device is called a *thermocouple,* and these devices find wide application in temperature measurement systems. This phenomenon was discovered in 1821 by Thomas Johann Seebeck.

The Seebeck effect is defined by

$$V = \int_{T_1}^{T_2} S_B(T) - S_A(T)dT$$

where V is the voltage created, S is the Seebeck coefficient, and T is the temperature at the junction.

The Thomson, Peltier, and Seebeck equations are related by

$$\prod = S \cdot T$$

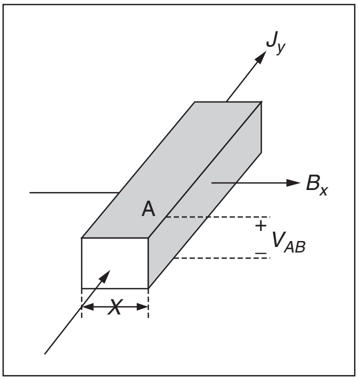


FIGURE 2-3 Hall-effect model.

Hall Effect. When a conductor carrying a current is inserted into a magnetic field that is perpendicular to the field, a force is exerted on the charged particles that constitute the current. The result is that the particles will be forced to the side of the conductor, leading to a buildup of positive charge on one side and negative charge on the other. This appears as a voltage across the conductor, given by

$$V_{AB} = -\frac{J_y B_x x}{en} = -vB_x x \tag{2-1}$$

where x is width of the conductor, B_x is magnetic field strength, J_y is current density, n is charge density, e is electronic charge, and v is velocity of charge flow.

This phenomenon is useful in the measurement of magnetic fields and in the determination of properties and characteristics of semiconductors, where the voltages are much larger than in conductors. See Fig. 2-3. This effect was discovered in 1879.

Faraday's Law of Induction. According to Faraday's law, in any closed linear path in space, when the magnetic flux ϕ (see Sec. 2.1.2) surrounded by the path varies with time, a voltage is induced around the path equal to the negative rate of change of the flux in webers per second.

$$V = -\frac{\partial \phi}{\partial t} \qquad \text{volts} \tag{2-2}$$

The minus sign denotes that the direction of the induced voltage is such as to produce a current opposing the flux. If the flux is changing at a constant rate, the voltage is numerically equal to the increase or decrease in webers in 1 s.

The closed linear path (or circuit) is the boundary of a surface and is a geometric line having length but infinitesimal thickness and not having branches in parallel. It can vary in shape or position.

If a loop of wire of negligible cross section occupies the same place and has the same motion as the path just considered, the voltage v will tend to drive a current of electricity around the wire, and this voltage can be measured by a galvanometer or voltmeter connected in the loop of wire. As with the path, the loop of wire is not to have branches in parallel; if it has, the problem of calculating the voltage shown by an instrument is more complicated and involves the resistances of the branches.

For accurate results, the simple Eq. (2-2) cannot be applied to metallic circuits having finite cross section. In some cases, the finite conductor can be considered as being divided into a large number of filaments connected in parallel, each having its own induced voltage and its own resistance. In other cases, such as the common ones of D.C. generators and motors and homopolar generators, where there are sliding and moving contacts between conductors of finite cross section, the induced voltage between neighboring points is to be calculated for various parts of the conductors. These can then be summed up or integrated. For methods of computing the induced voltage between two points, see text on electromagnetic theory.

In cases such as a D.C. machine or a homopolar generator, there may at all times be a conducting path for current to flow, and this may be called a *circuit,* but it is not a closed linear circuit without parallel branches and of infinitesimal cross section, and therefore, Eq. (2-2) does not strictly apply to such a circuit in its entirety, even though, approximately correct numerical results can sometimes be obtained.

If such a practical circuit or current path is made to enclose more magnetic flux by a process of connecting one parallel branch conductor in place of another, then such a change in enclosed flux does not correspond to a voltage according to Eq. (2-2). Although it is possible in some cases to describe a loop of wire having infinitesimal cross section and sliding contacts for which Eq. (2-2) gives correct numerical results, the equation is not reliable, without qualification, for cases of finite cross section and sliding contacts. It is advisable not to use equations involving $\partial\phi/\partial t$ directly on complete circuits where there are sliding or moving contacts.

Where there are no sliding or moving contacts, if a coil has N turns of wire in series closely wound together so that the cross section of the coil is negligible compared with the area enclosed by the coil, or if the flux is so confined within an iron core that it is enclosed by all N turns alike, the voltage induced in the coil is

$$V = -N\frac{\partial \phi}{\partial t} \qquad \text{volts} \tag{2-3}$$

In such a case, $N\phi$ is called the *number of interlinkages of lines of magnetic flux with the coil,* or simply, the *flux linkage*.

For the preceding equations, the change in flux may be due to relative motion between the coil and the magnetomotive force (mmf, the agent producing the flux), as in a rotating-field generator; it may be due to change in the reluctance of the magnetic circuit, as in an inductor-type alternator or microphone, variations in the primary current producing the flux, as in a transformer, variations in the current in the secondary coil itself, or due to change in shape or orientation of the loop of coil. For further study, refer to the Web site http://www.lectureonline.cl.msu.edu/~mmp/applist/induct/ faraday.htm.

2.1.2 Magnetic Fields

Early Concepts of Magnetic Poles. Substances now called *magnetic,* such as iron, were observed centuries ago as exhibiting forces on one another. From this beginning, the concept of magnetic poles evolved, and a quantitative theory built on the concept of these poles, or small regions of magnetic influence, was developed. André-Marie Ampère observed forces of a similar nature between conductors carrying currents. Further developments have shown that all theories of magnetic materials can be developed and explained through the magnetic effects produced by electric charge motions. Magnetic fields may be seen in Fig. 2-4.

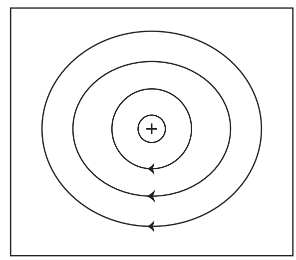


FIGURE 2-4 Magnetic fields.

Ampere's Formula. The magnetic field intensity dB produced at a point A by an element of a conductor ds (in meters) through which there is a current of i A is

$$dH = i\,ds\left(\frac{\sin\alpha}{4\pi r^2}\right) \qquad \text{A/m} \tag{2-4}$$

where r is the distance between the element ds and the point A, in meters, and α is the angle between the directions of ds and r. The intensity dH is perpendicular to the plane containing ds and r, and its direction is determined by the right-handed-screw rule given in Fig. 2-45.

The magnetic lines of force due to ds are concentric circles about the straight line in which ds lies. The field intensity produced at A by a closed circuit is obtained by integrating the expression for dH over the whole circuit.

An Indefinitely Long, Straight Conductor. The magnetic field due to an indefinitely long, straight conductor carrying a current of i A consists of concentric circles which lie in planes perpendicular to the axis of the conductor and have their centers on this axis. The magnetic field intensity at a distance of r m from the axis of the conductor is

$$H = \frac{i}{2\pi r} \qquad \text{A/m} \tag{2-5}$$

its direction being determined by the right-handed-screw rule (Sec. 2.1.22). See Fig. 2-5 for an illustration.

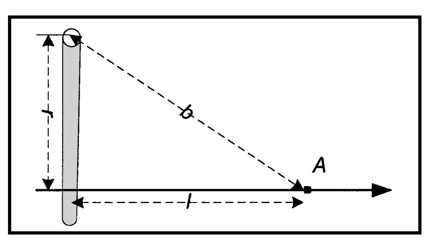


FIGURE 2-5 Magnetic field along the axis of a circular conductor.

Magnetic Field in Air Due to a Closed Circular Conductor. If the conductor carrying a current of i A is bent in the form of a ring of radius r m (Fig. 2-5), the magnetic field intensity at a point along the axis at a distance b m from the ring is

$$H = \frac{r^2 i}{2b^3} = \frac{r^2 i}{2(r^2 + l^2)^{3/2}} \qquad \text{A/m} \tag{2-6}$$

When $l = 0$,

$$H = \frac{i}{2r} \tag{2-7}$$

and when l is very great in comparison with r,

$$H = \frac{r^2 i}{2l^3} \tag{2-8}$$

Within a Solenoid. The magnetic field intensity within a solenoid made in the form of a *torus ring,* and also in the middle part of a *long, straight solenoid,* is approximately

$$H = n_1 i \qquad \text{A/m} \tag{2-9}$$

where i is the current in amperes and n_1 is the number of turns per meter length.

Magnetic Flux Density. The magnetic flux density resulting in free space, or in substances not possessing magnetic behaviors differing from those in free space, is

$$B = \mu H = 4\pi \times 10^{-7} H \tag{2-10}$$

where B is in teslas (or webers per square meter), H is in amperes per meter, and the constant $\mu_0 = 4\pi \times 10^{-7}$ is the *permeability* of free space and has units of henrys per meter. In the so-called practical system of units, the flux density is frequently expressed in *lines* or *maxwell per square inch*. The maxwell per square centimeter is called the *gauss*.

For substances such as iron and other materials possessing magnetic density effects greater than those of free space, a term μ_r is added to the relationship as

$$B = 4\pi \times 10^{-7} \mu_r H \tag{2-11}$$

where μ_r is the relative permeability of that substance under the conditions existing in it compared with that which would result in free space under the same magnetic-field-intensity condition. μ_r is a dimensionless quantity.

Magnetic Flux. The magnetic flux in any cross section of magnetic field is

$$\phi = \int B \cos \alpha dA \qquad \text{webers} \tag{2-12}$$

where α is the angle between the direction of the magnetic flux density B and the normal at each point to the surface over which A is measured. In the so-called practical system of units, the magnetic *line* (or *maxwell*) is frequently used, where 1 Wb is equivalent to 103 lines.

Density of Magnetic Energy. The magnetic energy stored per cubic meter of a magnetic field in free space is

$$\frac{dW}{dv} = \frac{1}{2}\mu_0 H^2 = 2\pi \times 10^{-7} H^2 = \frac{B^2}{2\mu_0} = \frac{B^2}{8\pi \times 10^{-7}} \qquad \text{J/m}^3 \tag{2-13}$$

In magnetic materials, the energy density stored in a magnetic field as a result of a change from a condition of flux density B_1 to that of B_2 can be expressed as

$$dW/dt = \int_{B_1}^{B_2} HdB \qquad (2\text{-}14)$$

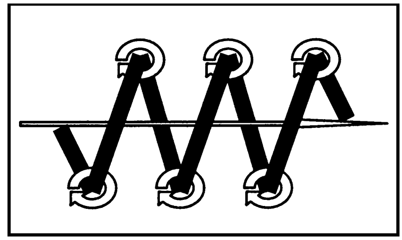


FIGURE 2-6 Magnetic field.

Flux Plotting. Flux plotting by a graphic process is useful for determining the properties of magnetic and other fields in air. The field of flux required is usually uniform along one dimension, and a cross section of it is drawn. The field is usually required between two essentially equal magnetic potential lines such as two iron surfaces. The field map consists of lines of force and equipotential lines which must intersect at right angles. For the graphic method, a field map of curvilinear squares is recommended when the problem is two dimensional. The squares are of different sizes, but the number of lines of force crossing every square is the same.

In sketching the field map, first draw those lines which can be drawn by symmetry. If parts of the two equipotential lines are straight and parallel to each other, the field map in the space between them will consist of lines which are practically straight, parallel, and equidistant. These can be drawn in. Then extend the series of curvilinear squares into other parts of the field, making sure, first, that all the angles are right angles and, second, that in each square the two diameters are equal, except in regions where the squares are evidently distorted, as near sharp comers of iron or regions occupied by current-carrying conductors. The diameters of a curvilinear square may be taken to be the distances between midpoints of opposite sides. An example of flux plotting may be seen in Fig. 2-6.

The magnetic field map near an iron comers is drawn as if the iron had a small fillet, that is, a line issues from an angle of 90° at 45° to the surface.

Inside a conductor which carries current, the magnetic field map is not made up of curvilinear squares, as in free space or air. In such cases, special rules for the spacing of the lines must be used. The equipotential lines converge to a point called the *kernel*.

Computer-based methods are now commonly available to do the detailed work, but the principles are unchanged.

2.1.3 Force Acting on Conductors

Force on a Conductor Carrying a Current in a Magnetic Field. Let a conductor of length l m carrying a current of i A be placed in a magnetic field, the density of which is B in teslas. The force tending to move the conductor across the field is

$$F = Bli \qquad \text{newtons} \qquad (2\text{-}15)$$

This formula presupposes that the direction of the axis of the conductor is at right angle to the direction of the field. If the directions of i and B form an angle α, the expression must be multiplied by sin α.

The force F is perpendicular to both i and B, and its direction is determined by the right-handed-screw rule. The effect of the magnetic field produced by the conductor itself is increase in the original flux density B on one side of the conductor and decrease on the other side. The conductor tends to move away from the denser field. A closed metallic circuit carrying current tends to move so as to enclose the greatest possible number of lines of magnetic force.

Force between Two Long, Straight Lines of Current. The force on a unit length of either of two long, straight, parallel conductors carrying currents of medium (that is, not near masses of iron) is

$$\frac{F}{L} = \frac{2 \times 10^{-7} i_1 i_2}{b} \qquad (2\text{-}16)$$

where F is in newtons and L (length of the long wires) and b (the spacing between them) are in the same units, such as meters.

The force is an attraction or a repulsion according to whether the two currents are flowing in the same or in opposite directions. If the currents are alternating, the force is pulsating. If i_1 and i_2 are effective values, as measured by A.C. ammeters, the maximum momentary value of the force may be as much as 100% greater than given by Eq. (2-16). The natural frequency (resonance) of mechanical vibration of the conductors may add still further to the maximum force, so a factor of safety should be used in connection with Eq. (2-16) for calculating stresses on bus bars.

If the conductors are straps, as is usual in bus bars, the following form of equation results for thin straps placed parallel to each other, b m apart:

$$\frac{F}{L} = \frac{2 \times 10^{-7} i_1 i_2}{s^2}\left(2s \tan^{-1}\frac{s}{b} - b\log_e \frac{s^2 + b^2}{b^2}\right) \qquad \text{N/m} \qquad (2\text{-}17)$$

where s is the dimension of the strap width in meters, and the thickness of the straps placed side by side is presumed small with respect to the distance b between them.

Pinch Effect. Mechanical force exerted between the magnetic flux and a current-carrying conductor is also present within the conductor itself and is called *pinch effect*. The force between the infinitesimal filaments of the conductor is an attraction, so a current in a conductor tends to contract the conductor. This effect is of importance in some types of electric furnaces where it limits the current that can be carried by a molten conductor. This stress also tends to elongate a liquid conductor.

2.1.4 Components, Properties, and Materials

Conductors, Semiconductors, and Insulators. An important property of a material used in electric circuits is its conductivity, which is a measure of its ability to conduct electricity. The definition of conductivity is

$$\sigma = J/E \qquad (2\text{-}18)$$

where J is current density, A/m^2, and E is electric field intensity, V/m.

The units of conductivity are thus the reciprocal of ohm-meter or siemens/meter. Typical values of conductivity for good conductors are 1000 to 6000 S/m. The reciprocal of conductivity is called *resistivity*. Section 4 gives extensive tabulations of the actual values for many different materials. Copper and aluminum are the materials usually used for distribution of electric energy and information. Semiconductors are a class of materials whose conductivity is in the range of 1 mS/m, though this number varies by orders of magnitude up and down. Semiconductors are produced by careful and precise modifications of pure crystals of germanium, silicon, gallium arsenide, and other materials. They form the basic building block for semiconductor diodes, transistors, silicon-controlled rectifiers, and integrated circuits. See Sec. 4.

Insulators (more accurately, dielectrics) are materials whose primary electrical function is to prevent current flow. These materials have conductivities of the order of nanosiemens/meter. Most insulating materials have nonlinear properties, being good insulators at sufficiently low electric field intensities and temperatures but breaking down at higher field strengths and temperatures. Figure 2-7 shows the energy levels of different materials. See Sec. 4 for extensive tabulations of insulating properties.

Gaseous Conduction. A gas is usually a good insulator until it is ionized, which means that electrons are removed from molecules. The electrons are then available for conduction. Ionized gases are good conductors. Ionization can occur through raising temperature, bringing the gas into contact with glowing metals, arcs, or flames, or by an electric current.

Electrolytes. In liquid chemical compounds known as *electrolytes,* the passage of an electric current is accompanied by a chemical change. Atoms of metals and hydrogen travel through the liquid in the direction of positive current, while oxygen and acid radicals travel in the direction of electron current. Electrolytic conduction is discussed fully in Sec. 24.

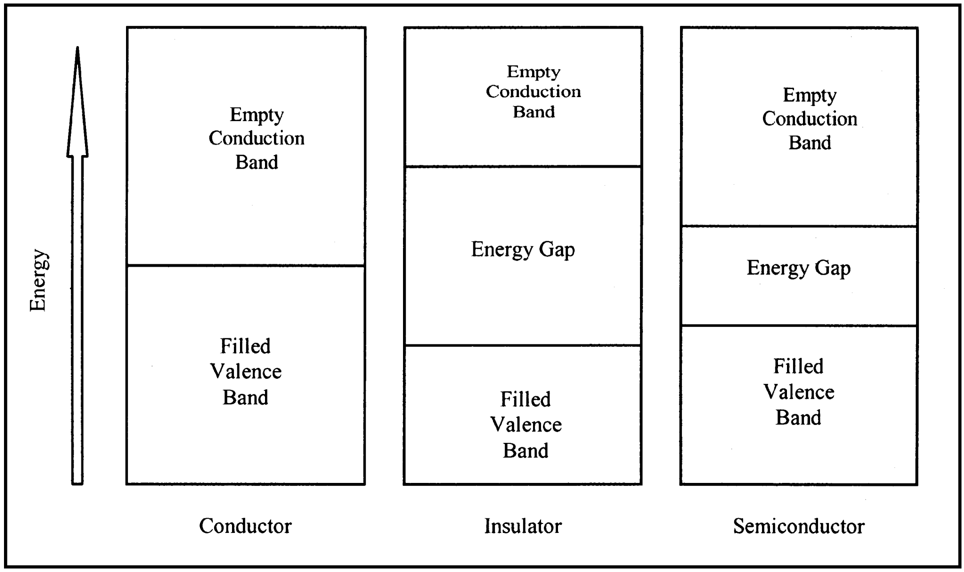


FIGURE 2-7 Component energy levels.

2.1.5 Resistors and Resistance

Resistors. A resistor is an electrical component or device designed explicitly to have a certain magnitude of resistance, expressed in ohms. Further, it must operate reliably in its environment, including electric field intensity, temperature, humidity, radiation, and other effects. Some resistors are designed explicitly to convert electric energy to heat energy. Others are used in control circuits, where they modify electric signals and energy to achieve desired effects. Examples include motor-starting resistors and the resistors used in electronic amplifiers to control the overall gain and other characteristics of the amplifier. A picture of a resistor may be seen in Fig. 2-8.

Ohm's Law. When the current in a conductor is steady and there are no voltages within the conductor, the value of the voltage v between the terminals of the conductor is proportional to the current i, or

$$v = ri \tag{2-19}$$

An example of Ohm's law may be seen in Fig. 2-9, where the coefficient of proportionality r is called the *resistance* of the conductor. The same law may be written in the form

$$i = gv \tag{2-20}$$

where the coefficient of proportionality $g = 1/r$ is called the *conductance* of the conductor. When the current is measured in amperes and the voltage in volts, the resistance r is in ohms and g is in siemens (often called mhos for reciprocal ohms). The phase of a resistor may be seen in Fig. 2-10.

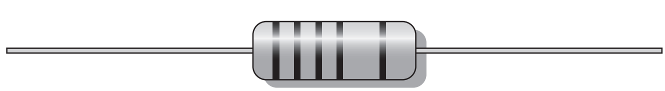


FIGURE 2-8 Resistor.

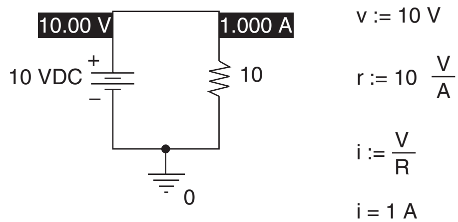


FIGURE 2-9 Ohm's law.

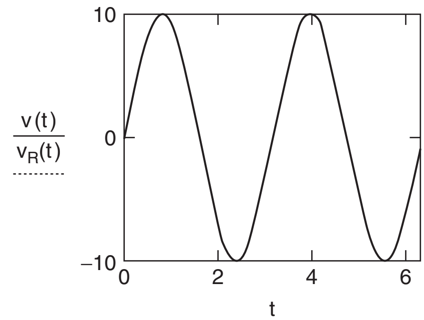


FIGURE 2-10 Phase of resistor.

Cylindrical Conductors. For current directed along the axis of the cylinder, the resistance r is proportional to the length l and inversely proportional to the cross section A, or

$$r = \rho \frac{l}{A} \tag{2-21}$$

where the coefficient of proportionality ρ (rho) is called the *resistivity* (or *specific resistance*) of the material. For numerical values of ρ for various materials, see Sec. 4.

The conductance of a cylindrical conductor is

$$g = \sigma \frac{A}{l} \tag{2-22}$$

where σ (sigma) is called the *conductivity* of the material. Since $g = 1/r$, the relation also holds that

$$\sigma = \frac{1}{\rho} \tag{2-23}$$

Changes of Resistance with Temperature. The resistance of a conductor varies with the temperature. The resistance of metals and most alloys increases with the temperature, while the resistance of carbon and electrolytes decreases with the temperature.

For usual conditions, as for about 100°C change in temperature, the resistance at a temperature t_2 is given by

$$R_{t_2} = R_{t_1}\left[1 + \alpha_{t_1}(t_2 - t_1)\right] \tag{2-24}$$

where R_{t_1} is the resistance at an initial temperature t, and α_{t_1} is called the *temperature coefficient of resistance* of the material for the initial temperature t_1. For copper having a conductivity of 100% of the International Annealed Copper Standard, $\alpha_{20} = 0.00393$, where temperatures are in degree Celsius (see Sec. 4).

An equation giving the same results as Eq. (2-24), for copper of 100% conductivity, is

$$\frac{R_{t_2}}{R_{t_2}} = \frac{234.4 + t_2}{234.4 + t_1} \tag{2-25}$$

where -234.4 is called the *inferred absolute zero* because if the relation held (which it does not over such a large range), the resistance at that temperature would be zero. For hard-drawn copper of 97.3% conductivity, the numerical constant in Eq. (2-25) is changed to 241.5. See Sec. 4 for values of these numerical constants for copper, and for other metals, see Sec. 4 under the metal being considered.

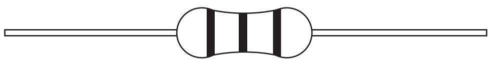


FIGURE 2-11 Inductor.

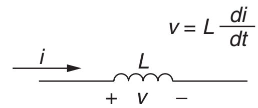


FIGURE 2-12 Inductor model and defining equation.

For 100% conductivity copper,

$$a_{t_1} = \frac{1}{234.4 + t_1} \tag{2-26}$$

When R_{t_1} and R_{t_2} have been measured, as at the beginning and end of a heat run, the "temperature rise by resistance" for 100% conductivity copper is given by

$$t_2 - t_1 = \frac{R_{t_2} - R_{t_1}}{R_{t_1}}(234.4 + t_1) \tag{2-27}$$

2.1.6 Inductors and Inductance

Inductors. An inductor is a circuit element whose behavior is described by the fact that it stores electromagnetic energy in its magnetic field. This feature gives it many interesting and valuable characteristics. In its most elementary form, an inductor is formed by winding a coil of wire, often copper, around a form that may or may not contain ferromagnetic materials. In this section, the behavior of the device at its terminal is discussed. Later, in the sections, on magnetic circuits, the device itself will be discussed. A picture of an inductor may be seen in Fig. 2-11.

Inductance. The property of the inductor that is useful in circuit analysis is called *inductance*. Inductance may be defined by either of the following equations:

$$v = L\frac{di}{dt} \qquad i = \frac{1}{L}\int_0^t v(\tau)d\tau + i(0) \tag{2-28}$$

or

$$W = \frac{1}{2}Li^2 \tag{2-29}$$

where L = coefficient of self-inductance
i = current through the coil of wire
v = voltage across the inductor terminals
W = energy stored in the magnetic field

Figure 2-12 shows the symbol for an inductor and the voltage-current relationship for the device. The unit of inductance is called the *henry* (H), in honor of American physicist Joseph Henry.

The phase of an inductor may be seen in Fig. 2-13.

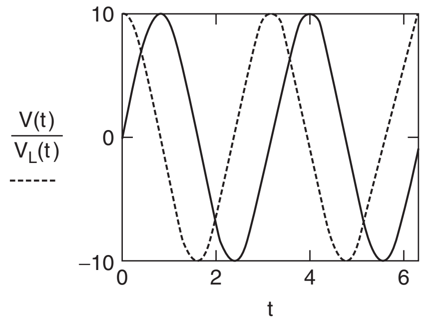


FIGURE 2-13 Phase of inductor.

Mutual Inductance. If two coils are wound on the same coil form, or if they exist in close proximity, then a changing current in one coil will induce a voltage in the second coil. This effect forms the basis for transformers, one of the most pervasive of all electrical

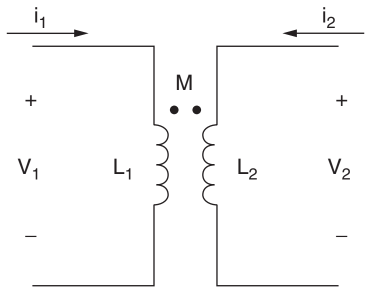


FIGURE 2-14 Mutual inductance model.

devices in use. Figure 2-14 shows the symbolic representation of a pair of coupled coils. The dots represent the direction of winding of the coils on the coil form in relation to the current and voltage reference directions. The equations become

$$v_1 = L_1\frac{di_1}{dt} + M\frac{di_2}{dt}$$
$$v_2 = M\frac{di_1}{dt} + L_2\frac{di_2}{dt} \tag{2-30}$$

Mutual inductance also can be a source of problems in electrical systems. One example is the problem, now largely solved, of cross talk from one telephone line to another.

2.1.7 Capacitors and Capacitance

Charge Storage. A capacitor is a circuit element that is described through its principal function, which is to store electric energy. This property is called *capacitance*. In its simplest form, a capacitor is built with two conducting plates separated by a dielectric. A picture of a conductor may be seen in Fig. 2-15. Figure 2-16 shows the two usual symbols for a capacitor and the defining directions for voltage and current. These equations further describe the capacitor.

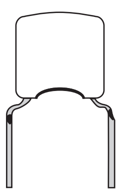


FIGURE 2-15 Capacitor.

$$i = C\frac{dv}{dt} \quad v(t) = \frac{1}{C}\int_0^t i(\tau)d\tau + v(0) \tag{2-31}$$

or

$$W = \frac{1}{2}Cv^2 \tag{2-32}$$

where W = energy stored in the capacitor
τ = dummy variable representing time
C = capacitance in farads

The unit of capacitance is the farad (F), named in honor of English physicist Michael Faraday.
The phase of a capacitor may be seen in Fig. 2-17.

2.1.8 Power and Energy

Power. The power delivered by an electrical source to an electrical device is given by

$$p(t) = v(t)i(t) \tag{2-33}$$

where p = power delivered
v = voltage across the device
i = current delivered to the device

The choice of algebraic sign is important. See Fig. 2-18*a*.

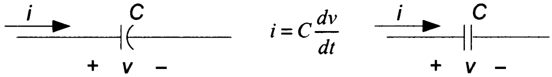


FIGURE 2-16 Capacitor-two symbols and defining equation.

If the device is a resistor, then the power delivered to the device is

$$p(t) = v(t)i(t) = i^2(t)R = \frac{v^2(t)}{R} \qquad (2\text{-}34)$$

an equation known as *Joule's law*. In SI units, the unit of power is the watt (W), in honor of eighteenth century Scottish engineer James Watt.

Energy. The energy delivered by an electrical source to an electrical device is given by

$$W = \int_{t_1}^{t_2} v(t)i(t)dt \qquad (2\text{-}35)$$

where the times t_1 and t_2 represent the starting and ending times of the energy delivery. In SI units, the unit of energy is the joule (J), in honor of English physicist James Joule. Power and energy are also related by the equation

$$p(t) = \frac{dW(t)}{dt} \qquad (2\text{-}36)$$

A commonly used unit for electric energy measurement is a kilowatthour (kWh), which is equal to 3.6×10^6 joules.

Energy Density and Power Density. At times it is useful to evaluate materials and media by comparing their energy storage capability on a unit volume basis. The SI unit is joules per cubic meter, though conversion to other convenient combinations of units is possible. Power density is often an important consideration in, for example, heat or energy flow. The SI unit is watts per square meter, although any convenient unit system can be used.

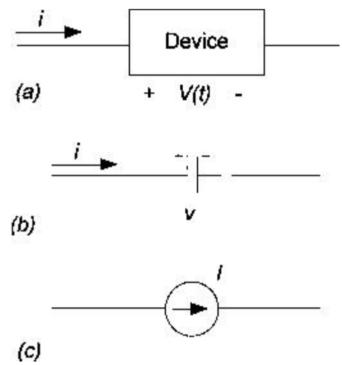


FIGURE 2-18 (*a*) Electrical device with definitions of voltage and current directions; (*b*) constant (D.C.) voltage source; (*c*) constant (D.C.) current source.

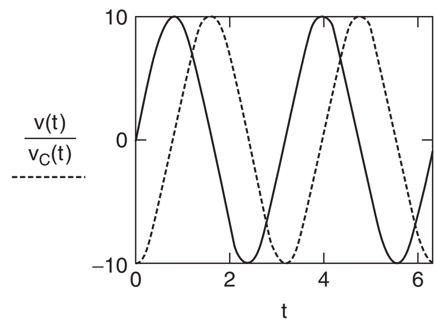


FIGURE 2-17 Phase of capacitor.

2.1.9 Physical Laws for Electric and Magnetic Circuits

Maxwell's Equations. Throughout much of the nineteenth century, engineers and physicists developed the theories that describe electricity and magnetism and their interrelations. In contemporary vector calculus notation, four equations can be written to describe the basic theory of electromagnetic fields. Collectively, they are known as *Maxwell's equations*, recognizing the work of James Clerk Maxwell's, a Scottish physicist, who solidified the theory. (Some writers consider only the first two as Maxwell's equations, calling the last two as supplementary equations.) The following symbols will be used in the description of Maxwell's equations:

E electric field intensity	*vol* (or V) enclosed volume in space
D electric flux density	L length of boundary around a surface
H magnetic field intensity	ρ electric charge density per unit volume
B magnetic flux density	J electric current density

Faraday's Law. Faraday observed that a time-varying magnetic field develops a voltage that can be observed and measured. This law is the basis for inductors. One common form of expressing the law is the equation

$$v = -\frac{\partial \phi}{\partial t}$$

where v is the voltage induced by the changing flux. The negative sign expresses the principle of conservation of energy, indicating that the direction of the voltage is such as to oppose the changing flux. This effect is known as *Lenz's law*.

In vector calculus notation, Faraday's law can be written in both integral and differential form. In integral form, the equation is

$$\oint EdL = \int_s \frac{\partial B}{\partial t} dS$$

where the line integral completely encircles the surface over which the surface integral is taken. In differential (point) form, Faraday's law becomes

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{2-37}$$

Ampere's Law. French physicist André-Marie Ampère developed the relation between magnetic field intensity and electric current that is a dual of Faraday's law. The current consists of two components, a steady or constant component and a time-varying component usually called *displacement current*. In vector calculus notation, Ampere's law is written first in integral form and then in differential form:

$$\oint HdL = I + \int_s \frac{dD}{dt} dS \tag{2-38}$$

$$\nabla \times H = J + \frac{\partial D}{\partial t} \tag{2-39}$$

An illustration of Ampere's law may be seen in Fig. 2-19. For more information, please refer to the Web site http://www.ee.byu.edu./em/amplaw2.htm.

Gauss's Law. Carl F. Gauss, a German physicist, stated the principle that the displacement current flowing over the surface of a region (volume) in space is equal to the charge enclosed. In integral and differential form, respectively, this law is written

$$\oint_s DdS = \int_{\text{vol}} \rho dv \tag{2-40}$$

$$\nabla \cdot D = \rho \tag{2-41}$$

For further study, please refer to the Web site http://www.ee.byu.edu/ee/em/eleclaw.htm.

An illustration of Gauss's law may be seen in Fig. 2-20.

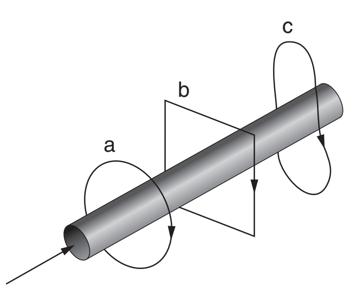


FIGURE 2-19 Ampere's law illustration.

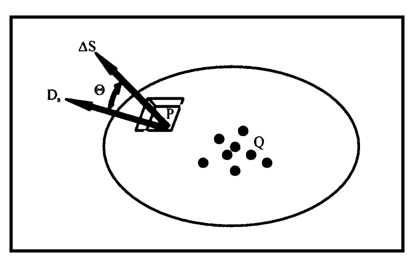


FIGURE 2-20 Gauss's law illustration.

Gauss's Law for Magnetics. One of the postulates of electromagnetism is that there are no free magnetic charges but these charges always exist in pairs. While searches are continually being made, and some claims of discovery of free charges have been made, the postulate is still adequate to explain observations in cases of interest here. A consequence of this postulate is that, for magnetics, Gauss's law become

$$\oint_s BdS = 0 \tag{2-42}$$

$$\nabla \cdot B = 0 \tag{2-43}$$

Kirchhoff's Laws. In the analysis and design of electric circuits, a fundamental principle implies that the dimensions are small. This means that it is possible to neglect the spatial variations in electromagnetic quantities. Another way of saying this is that the dimensions of the circuit are small compared with the wavelengths of the electromagnetic quantities and thus that it is necessary to consider only time variations. This means that Maxwell's equations, which are partial integrodifferential equations, become ordinary integrodifferential equations in which the independent variable is time, represented by t.

Kirchhoff's Current Law. The assumption of small dimensions means that no free electric charges can exist in the region in which a circuit is being analyzed. Thus, Gauss's law (in integral form) becomes

$$\sum i = 0 \tag{2-44}$$

at any point in the circuit. The points of interest usually will be *nodes,* points at which three or more wires connect circuit elements together. This law will be abbreviated KCL and was enunciated by German physicist, Gustav Robert Kirchhoff. It is one of the two fundamental principles of circuit analysis. Figure 2-21 shows a sample circuit simulated in PSPICE. We can see the current flowing through the 3 Ω resistor (3.5 A) is equal to the sum of the current flowing through the 6 Ω resistor (1.5 A) and the current flowing through the 1.5 Ω resistor (2 A).

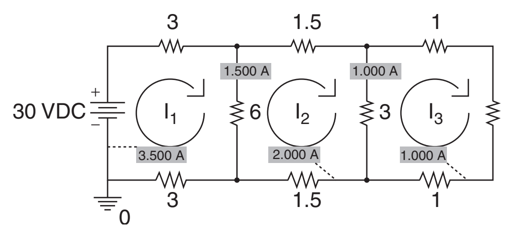


FIGURE 2-21 Kirchhoff's current law.

Kirchhoff's Voltage Law. The second fundamental principle, abbreviated KVL, follows from applying the assumption of small size to Faraday's law in integral form. Since the circuit is small, it is possible to take the surface integral of magnetic flux density as zero and then to state that the sum of voltages around any closed path is zero. In equation form, it can be written as

$$\sum v = 0 \tag{2-45}$$

Figure 2-22 shows a sample circuit simulated in PSPICE. We have

$$\sum V = V_{0->V_1} + V_{V_1->V_2} + V_{V_2->V_7} + V_{V_7->0} = -30 + 10.5 + 9 + 10.5 = 0$$

2.1.10 Electric Energy Sources and Representations

Sources. In circuit analysis, the goal is to start with a connected set of circuit elements such as resistors, capacitors, operational amplifiers, and other devices, and to find the voltages across and currents through each element, additional quantities, such as power dissipated, are often computed. To energize the circuit, sources of electric energy must be connected. Sources are modeled in various ways.

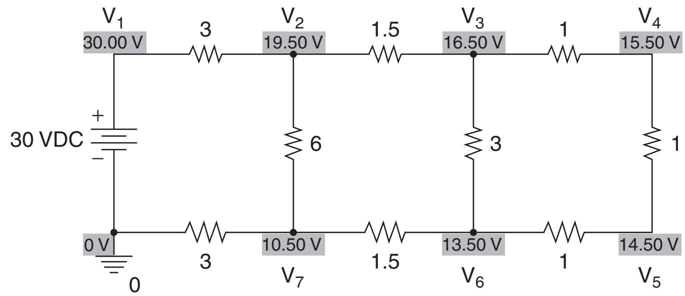


FIGURE 2-22 Kirchhoff's voltage law.

One convenient classification is to consider constant (dc) sources, sinusoidal (ac) sources, and general time-varying sources. The first two are of interest in this section.

DC Sources. Some sources, such as batteries, deliver electric energy at a nearly constant voltage, and thus they are modeled as constant voltage sources. The term *dc sources* basically means *direct-current sources,* but it has come to stand for constant sources as well. Figure 2-23 shows the standard symbol for a dc source. Other sources are modeled as dc current (or constant-current) sources. Figure 2-18*b* and *c* show the symbols used for these models.

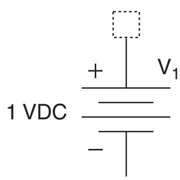


FIGURE 2-23 D.C. source.

AC Sources. Most of the electric energy used in the world is generated, distributed, and utilized in sinusoidal form. Thus, beginning with Charles P. Steinmetz, a German-American electrical engineer, much effort has been devoted to finding efficient ways to analyze and design circuits that operate under sinusoidal excitation conditions. Sources of this type are frequently called ac (for alternating current) sources. Figure 2-24 shows the standard symbol for an ac source. The most general expression for a voltage in sinusoidal form is of the type

$$v(t) = V_m \cos(2\pi ft + \alpha) = V_m \cos(\omega t + \alpha) \tag{2-46}$$

and, for a current

$$i(t) = I_m \cos(2\pi ft + \beta) = I_m \cos(\omega t + \beta) \tag{2-47}$$

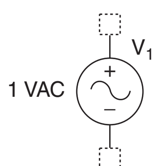


FIGURE 2-24 AC source.

Some writers use sine functions instead of cosine functions, but this has only the effect of changing the angles α and β.

These expressions have three identifying characteristics, the maximum or peak value (V_m or I_m), the phase angle (α or β), and the frequency [f, measured in hertz (Hz) or cycles per second, or ω, measured in radians/second]. A powerful method of circuit analysis depends on these observations. It is called *phasor analysis*.

2.1.11 Phasor Analysis

The Imaginary Operator. A term that arises frequently in phasor analysis is the imaginary operator

$$j = \sqrt{-1} \tag{2-48}$$

(Electrical engineers use j, since i is reserved as the symbol for current. Mathematicians, physicists, and others are more likely to use i for the imaginary operator.)

Euler's Relation. A relationship between trigonometric and exponential functions, known as *Euler's relation*, plays an important role in phasor analysis. The equation is

$$e^{jx} = \cos x + j\sin x \tag{2-49}$$

If this equation is solved for the trigonometic terms, the result is

$$\cos x = \frac{e^{jx} + e^{-jx}}{2} \tag{2-50}$$

$$\sin x = \frac{e^{jx} - e^{-jx}}{2} \tag{2-51}$$

In phasor analysis, this equation is used by writing it as

$$e^{j(wt+\alpha)} = e^{jwt}e^{j\alpha} = \cos(wt + \alpha) + j\sin(wt + \alpha) \tag{2-52}$$

Thus, it is observed that the cosine term in the preceding expressions for voltage and current is equal to the real-part term from Euler's relation. Thus, it will be seen possible to substitute the general exponential term for the cosine term in the source expressions, then, to find the solution (currents and voltages) to the exponential excitation, and finally, to take the real part of the result to get the final answer.

Steady-State Solutions. When the complete solution for current and voltage in a linear, stable, time-invariant circuit is found, two types of terms are found. One type of term, called the *complementary function* or *transient solution,* depends only on the elements in the circuit and the initial energy stored in the circuit when the forcing function is connected. If the circuit is stable, this term typically becomes very small in a short time.

The second type of term, called the *particular integral* or *steady-state solution,* depends on the circuit elements and configuration and also on the forcing function. If the forcing function is a single-frequency sinusoidal function, then it can be shown that the steady-state solution will contain terms at this same frequency but with differing amplitudes and phases. The goal of phasor analysis is to find the amplitudes and phases of the voltages and currents in the solution as efficiently as possible, since the frequency is known to be the same as the frequency of the forcing function.

Definition of a Phasor. The phasor representation of a sinusoidal function is defined as a complexnumber containing the amplitude and phase angle of the original function. Specifically, if

$$v(t) = V_m\cos(wt + \alpha) = V_m\sin(wt + \alpha + \pi/2) \tag{2-53}$$

then the phasor representation is given by

$$V = V_m e^{j\alpha} \tag{2-54}$$

A phasor can be converted to a sinusoidal time function by using the definition in reverse. See Sec. 2.1.12 for an alternative definition of a phasor, which differs only by a multiplicative constant.

Phasor Algebra. It is necessary at times to perform arithmetic and algebraic operations on phasors. The rules of phasor algebra are identical with those of complex number algebra and vector algebra. Specifically,

$$\begin{aligned} V_1e^{j\alpha_1} \pm V_2e^{j\alpha_2} &= (V_1\cos\alpha_1 \pm V_2\cos\alpha_2) + j(V_1\sin\alpha_1 \pm V_2\sin\alpha_2) \\ (V_1e^{j\alpha_1})(V_2e^{j\alpha_2}) &= (V_1V_2)\, e^{j(\alpha_1+\alpha_2)} \\ \frac{V_1e^{j\alpha_1}}{V_2e^{j\alpha_2}} &= \frac{V_1}{V_2}e^{j(\alpha_1-\alpha_2)} \end{aligned} \tag{2-55}$$

A few examples will show the calculations. In the examples, the angles are expressed in radians. Sometimes degrees are used instead, at the option of the analyst.

$$5e^{j\pi/5} + 4e^{j2\pi/3} = (4.05 + j2.94) + (-2.00 + j3.46) = 6.72e^{j1.26}$$
$$(8e^{j\pi/4})(1.3e^{j3\pi/5}) = 10.4e^{j2.67} = -9.27 + j4.72 \tag{2-56}$$

With a modern electronic calculator or one of many suitable computer programs, it is possible to perform these calculations readily, though they may appear tedious.

Integration and Differentiation Operations. Let a phasor be represented by

$$P_1 = P_m e^{j\omega t} e^{j\theta} \tag{2-57}$$

where the frequency is included for completeness. Differentiation and integration become, respectively,

$$\frac{dP_1}{dt} = j\omega P_m e^{j\omega t} e^{j\theta} = \omega P_m e^{j\omega t} e^{j(\theta+90^\circ)} \tag{2-58}$$

$$\int P_1 dt = \frac{1}{j\omega} P_1 e^{j\omega t} e^{j\theta} = \frac{1}{\omega} P_1 e^{j\omega t} e^{j(\theta-90^\circ)} \tag{2-59}$$

Reactance and Susceptance. For an inductor, the ratio of the phasor voltage to the phasor current is given by *jwL*. This quantity is called the *reactance* of the inductor, and its reciprocal is called *susceptance*. For a capacitor, the ratio of phasor voltage to phasor current is $1/(j\omega C) = -j(\omega C)$. This quantity is called the *reactance* of a capacitor. Its reciprocal is called *susceptance*. The usual symbol for reactance is *X,* and for susceptance, *B*.

Impedance and Admittance. Analysis of ac circuits requires the analyst to replace each inductor and capacitor with appropriate susceptances or reactances. Resistors and constant controlled sources are unchanged. Application of any of the methods of circuit analysis will lead to a ratio of a voltage phasor to a current phasor. This ratio is called *impedance*. It has a real (or resistive) part and an imaginary (or reactive) part. Its reciprocal is called *admittance*. The real part of admittance is called the *conductive* part, and the imaginary part is called the *susceptive part*. In Sec. 2.1.15, an analysis of a circuit shows the use of these ideas.

2.1.12 AC Power and Energy Considerations

Effective or RMS Values. If a sinusoidal current $i(t) = I_m \cos(\omega t + \alpha)$ flows through a resistor of $R\ \Omega$, then, over an integral number of cycles, the average power delivered to the resistor is found to be

$$P_{\text{avg}} = \frac{I_m^2}{2} R \tag{2-60}$$

This amount of power is identical to the amount of power that would be delivered by a constant (dc) current of $I_m/\sqrt{2}$ amperes. Thus, the effective value of an ac current (or voltage) is equal to the maximum value divided by $\sqrt{2}$.

The effective value is commonly used to describe the requirements of ac systems. For example, in North America, rating a light bulb at 120 V implies that the bulb should be used in a system where the effective voltage is 120 V. In turn, the voltages and currents quoted for distribution systems are effective values.

An alternative term is root-mean-square (rms) value. This term follows from the formal definition of effective or rms values of a function,

$$F_{\text{rms}} = \sqrt{\frac{1}{T}\int_{t_0}^{t_0+T} (f(t))^2 dt} \tag{2-61}$$

Frequently, when phasor ideas are being used, effective rather than peak values are implied. This is quite common in electric power system calculations, and it is necessary for the engineer simply to determine which is being used and to be consistent.

Power Factor. When the voltage across a device and the current through a device are given, respectively, by

$$v(t) = V_m \cos(\omega t + \alpha) \tag{2-62}$$

and

$$i(t) = I_m \cos(\omega t + \beta) \tag{2-63}$$

a computation of the average power over an integral number of cycles gives

$$P_{\text{avg}} = \frac{V_m I_m}{2} \cos(\alpha - \beta) \tag{2-64}$$

and

$$P_{\text{avg}} = V_{\text{eff}} I_{\text{eff}} \cos(\alpha - \beta) \tag{2-65}$$

The angle $(\alpha - \beta)$, which is the phase difference between the voltage and current, is called the *power factor angle*. The cosine of the angle is called the *power factor* because it represents the ratio of the average power delivered to the product of voltage and current.

Reactive Voltamperes. When the voltage across a device and the current through a device are given, respectively, by

$$v(t) = V_m \cos(\omega t + \alpha) \tag{2-66}$$

and

$$i(t) = I_m \cos(\omega t + \beta) \tag{2-67}$$

a computation of the power delivered to the device as a function of time shows

$$p(t) = \frac{V_m I_m}{2} [\cos(\alpha - \beta) + \cos(2\omega t + \alpha + \beta)] \tag{2-68}$$

In addition to the constant term that represents the average power, there is a double-frequency term that represents energy that is interchanged between the electric and magnetic fields of the device and the source. This quantity is called by the term *reactive voltamperes* (vars). It may be shown that

$$\text{var} = \frac{V_m I_m}{2} \sin(\alpha - \beta) \tag{2-69}$$

and

$$\text{var} = V_{\text{eff}} I_{\text{eff}} \sin(\alpha - \beta) \tag{2-70}$$

Power and Vars. If the phasor voltage across a device and the phasor current through the device are given, respectively, by

$$V_1 = V_{\text{eff}} e^{j\alpha} \tag{2-71}$$

and

$$I_2 = I_{\text{eff}}\, e^{j\beta} \tag{2-72}$$

then the expression

$$VA = V_{\text{eff}} I^*_{\text{eff}} = V_{\text{eff}} I_{\text{eff}}[\cos(\alpha - \beta) + j\sin(\alpha - \beta)] \tag{2-73}$$

where * represents the complex conjugate, which may be used to find both average power and vars. The real part of the expression is the average power, while the imaginary part is the vars.

2.1.13 Controlled Sources

Models. When circuits containing electronic devices such as amplifiers and similar devices are analyzed or designed, it is necessary to have a linear circuit model for the electronic device. These devices have a minimum of three terminals. Currents flow between terminal pairs, and voltages appear across terminal pairs. One terminal may not be common to both pairs. A useful model is provided by a controlled source. Four such models may be distinguished, as shown in Fig. 2-25. Examples of use will appear in the paragraphs on circuit analysis.

Voltage-Controlled Voltage Source (VCVS). If the voltage across one terminal pair is proportional to the voltage across a second terminal pair, then the model of Fig. 2-26*a* may be used. In this model, the output voltage v_{xz} is proportional to the input voltage v_{yx}, with a proportionality constant A. It should be noted, however, that this device is not usually reciprocal, that is, impression of a voltage at the terminals xz will not lead to a voltage at terminals yz'.

Voltage-Controlled Current Source. If the current flow between a terminal pair is proportional to the voltage across another pair, then the appropriate model is a voltage-controlled current source (VCCS). See Fig. 2-26*b*.

Current-Controlled Current Source. If the current flow between a terminal pair is proportional to the current through another terminal pair, then the appropriate model is a current-controlled current source (CCCS), as shown in Fig. 2-26*c*.

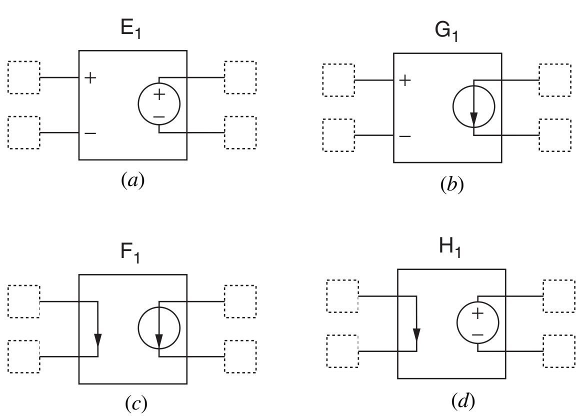


FIGURE 2-25 (*a*) Voltage-controlled voltage source; (*b*) Voltage-controlled current source; (*c*) Current-controlled current source; (*d*) Current-controlled voltage source.

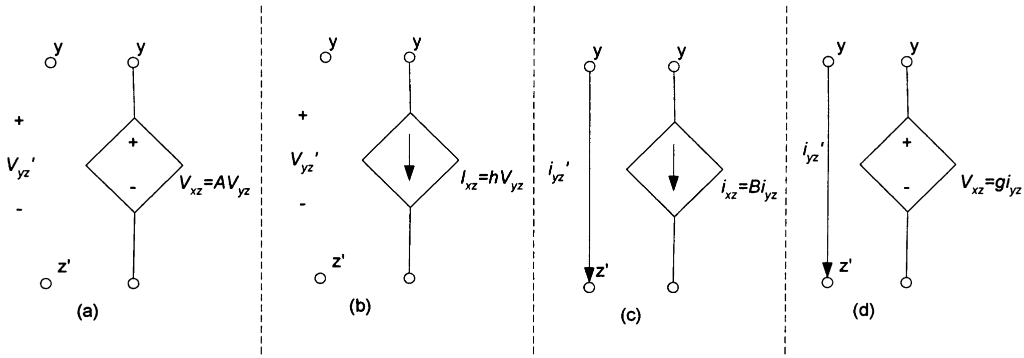


FIGURE 2-26 Controlled sources: (*a*) voltage-controlled voltage source; (*b*) voltage-controlled current source; (*c*) current-controlled current source; (*d*) current-controlled voltage source.

Current-Controlled Voltage Source. If the voltage across one terminal pair is proportional to the current flow through another terminal pair, then the appropriate model is a current-controlled voltage source (CCVS), as shown in Fig. 2-26*d*.

2.1.14 Methods for Circuit Analysis

Circuit Reduction Techniques. When a circuit analyst wishes to find the current through or the voltage across one of the elements that make up a circuit, as opposed to a complete analysis, it is often desirable to systematically replace elements in a way that leaves the target elements unchanged, but simplifies the remainder in a variety of ways. The most common techniques include series/parallel combinations, wye/delta (or tee/pi) combinations, and the Thevenin-Norton theorem.

Series Elements. Two or more electrical elements that carry the same current are defined as being in series. Figure 2-27 shows a variety of equivalents for elements connected in series.

Parallel Elements. Two or more electrical elements that are connected across the same voltage are defined as being in parallel. Figure 2-28 shows a variety of equivalents for circuit elements connected in parallel.

Wye-Delta Connections. A set of three elements may be connected either as a wye, shown in Fig. 2-29*a*, or a delta, shown in Fig. 2-29*b*. These are also called *tee* and *pi connections*, respectively. The equations give equivalents, in terms of resistors, for converting between these connection forms.

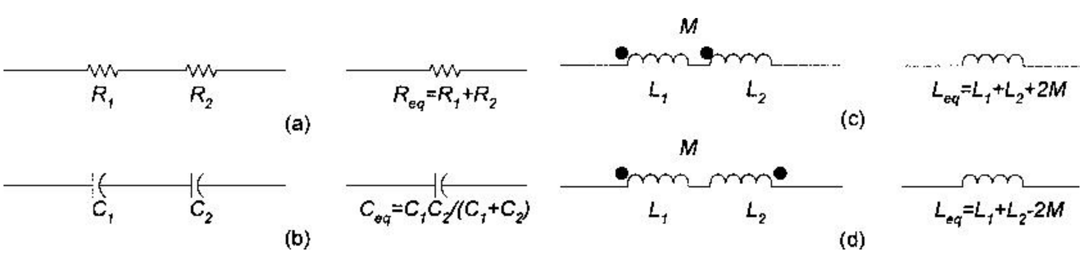


FIGURE 2-27 Series-connected elements and equivalents: (*a*) resistors in series; (*b*) capacitors in series; (*c*) inductors in series, aiding fluxes; (*d*) inductors in series, opposing fluxes.

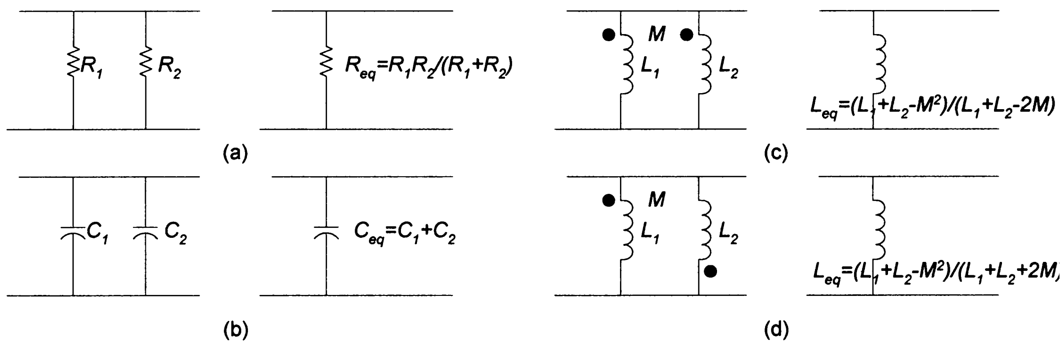


FIGURE 2-28 Parallel-connected elements and equivalents: (*a*) resistors in parallel; (*b*) capacitors in parallel; (*c*) inductors in parallel, aiding fluxes; (*d*) inductors in parallel, opposing fluxes.

$$R_c = \frac{R_1R_2 + R_1R_3 + R_2R_3}{R_1}$$
$$R_b = \frac{R_1R_2 + R_1R_3 + R_2R_3}{R_2} \qquad (2\text{-}74)$$
$$R_a = \frac{R_1R_2 + R_1R_3 + R_2R_3}{R_3}$$

$$R_1 = \frac{R_aR_b}{R_a + R_b + R_c}$$
$$R_2 = \frac{R_aR_c}{R_a + R_b + R_c} \qquad (2\text{-}75)$$
$$R_3 = \frac{R_bR_c}{R_a + R_b + R_c}$$

In practice, application of one of these conversion pairs will lead to additional series or parallel combinations that can be further simplified.

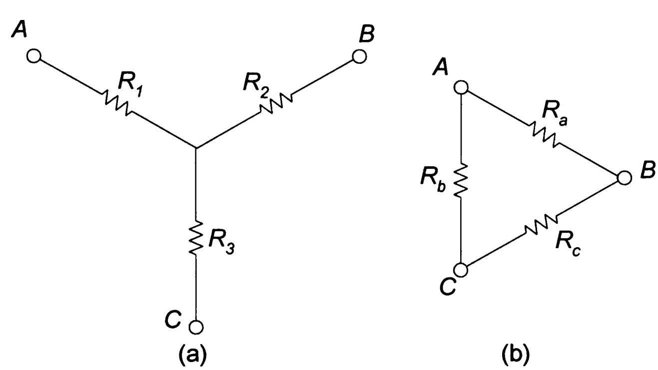


FIGURE 2-29 (*a*) Wye-connected elements; (*b*) delta-connected elements.

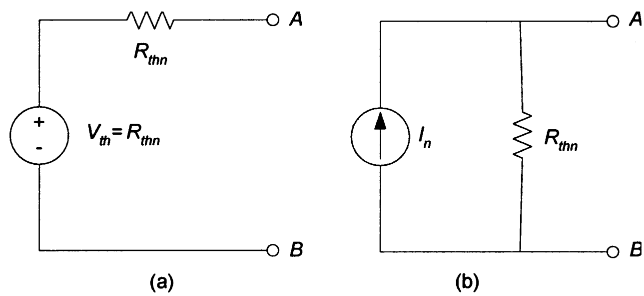


FIGURE 2-30 (*a*) Thevenin equivalent circuit model; (*b*) Norton equivalent ciruit model.

Thevenin-Norton Theorem. The Thevenin theorem and its dual, the Norton theorem, provide the engineer with a convenient way of characterizing a network at a terminal pair. The method is most useful when one is considering various loads connected to a pair of output terminals. The equivalent can be determined analytically, and in some cases, experimentally. Terms used in these paragraphs are defined in Fig. 2-30.

Thevenin Theorem. This theorem states that at a terminal pair, any linear network can be replaced by a voltage source in series with a resistance (or impedance). It is possible to show that the voltage is equal to the voltage at the terminal pair when the external load is removed (open circuited), and that the resistance is equal to the resistance calculated or measured at the terminal pair with all independent sources de-energized. De-energization of an independent source means that the source voltage or current is set to zero but that the source resistance (impedance) is unchanged. Controlled (or dependent) sources are not changed or de-energized.

Norton Theorem. This theorem states that at a terminal pair, any linear network can be replaced by a current source in parallel with a resistance (or impedance). It is possible to show that the current is equal to the current that flows through the short-circuited, terminal pair when the external load is short circuited, and that the resistance is equal to the resistance calculated or measured at the terminal pair with all independent sources de-energized. De-energization of an independent source means that the source voltage or current is set to zero but that the source resistance (impedance) is unchanged. Controlled (or dependent) sources are not changed or de-energized.

Thevenin-Norton Comparison. If the Thevenin equivalent of a circuit is known, then it is possible to find the Norton equivalent by using the equation

$$V_{th} = I_n R_{thn} \tag{2-76}$$

as indicated in Fig. 2-30.

Thevenin-Norton Example. Figure 2-31*a* shows a linear circuit with a current source and a voltage-controlled voltage source. Figure 2-31*b* shows a calculation of the Thevenin or open-circuit voltage. Figure 2-31*c* shows a calculation of the Norton or short-circuit current. Figure 2-31*d* shows the final Norton and Thevenin equivalent circuits.

2.1.15 General Circuit Analysis Methods

Node and Loop Analysis. Suppose b elements or branches are interconnected to form a circuit. A complete solution for the network is one that determines the voltage across and the current through each element. Thus, $2b$ equations are needed. Of these, b are given by the voltage-current relations, for example, Ohm's law, for each element. The others are obtained from systematic application of Kirchhoff's voltage and current laws.

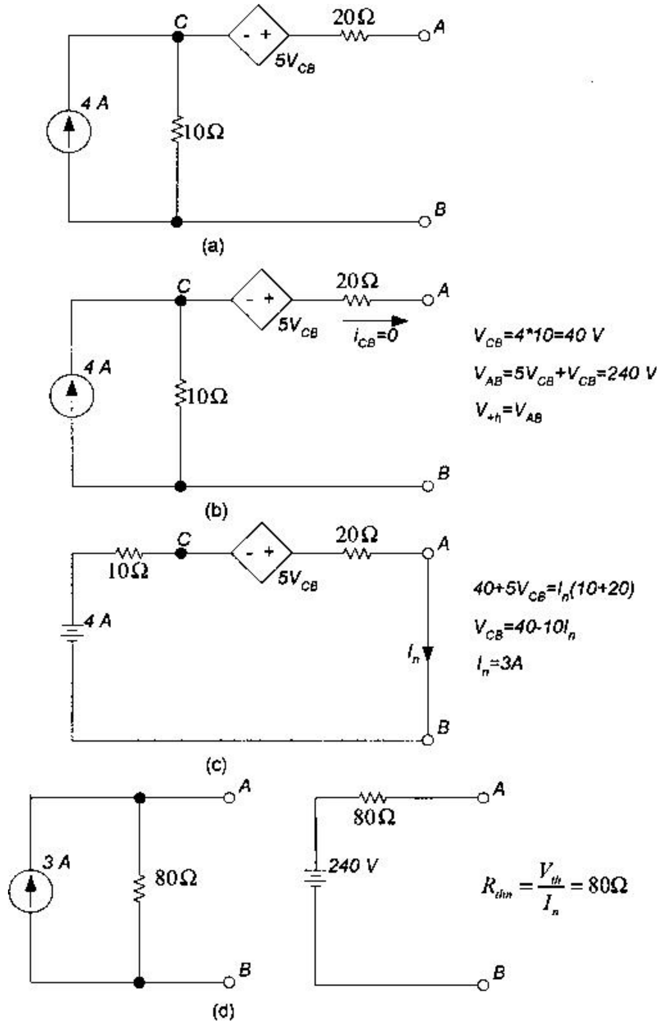


FIGURE 2-31 To illustrate Thevenin-Norton theorem: (*a*) example circuit; (*b*) calculation of Thevenin voltage; (*c*) calculation of Norton current; (*d*) Norton and Thevenin equivalent circuits.

Define a point at which three or more elements or branches are connected as a node (some writers call this an *essential node*). Suppose that the circuit has n such nodes or points. It is possible to write Kirchhoff's current law equations at each node, but one will be redundant, that is, it can be derived from the others. Thus, n–1 KCL equations can be written, and these are independent.

This means that to complete the analysis, it is necessary to write $[b-(n-1)]$ KVL equations, and this is possible, though care must be taken to ensure that they are independent. In practice, it is typical that either KCL or KVL equations are written, but not both. Sufficient information is usually available from either set. Which set is chosen depends on the analyst, the comparative number of equations, and similar factors.

Nodal Analysis. Figure 2-32*a* shows a typical node in a circuit that is isolated for attention. The voltage on this node is measured or calculated with a reference somewhere in the circuit, often but not always the node that is omitted in the analysis. Other nodes, including the reference node, are shown along with connecting elements. To illustrate the technique, five additional nodes are chosen, including the reference nodes. The boxes labeled *Y* are called *admittances*.

Kirchhoff's current law written at the node states that

$$i_{k-2} + i_{k-1} + i_k + i_{k+1} + i_{k+2} = I_{in,k} \tag{2-77}$$

An expression for each of the currents can be written

$$i_{k+1} = (V_k - V_{k+1})Y_{k+1} \tag{2-78}$$

When all the equations that can be written are written, collected, and organized into matrix format, the general result is

$$\begin{bmatrix} Y_{11} + Y_{12} + Y_{13} + \cdots & -Y_{12} & -Y_{13} & \cdots \\ -Y_{21} & Y_{21} + Y_{22} + Y_{23} + \cdots & -Y_{23} & \cdots \\ -Y_{31} & -Y_{32} & Y_{31} + Y_{32} + Y_{33} + \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ \vdots \end{bmatrix} \tag{2-79}$$

where the square matrix describes the circuit completely, the column matrix (vector) of voltages describes the dependent variables which are the node voltages, and the column matrix of currents describes the forcing function currents that enter each node.

Nodal Analysis with Controlled Sources. If a controlled source is present, it is most convenient to use the Thevenin-Norton theorem to convert the controlled source to a voltage-controlled current

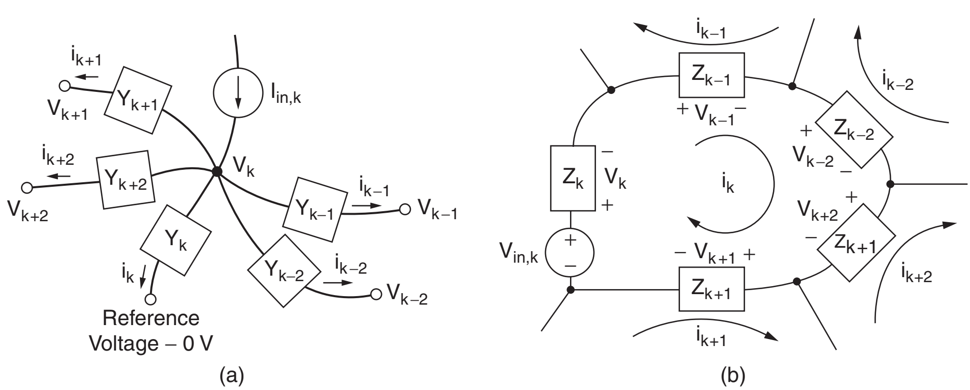


FIGURE 2-32 To illustrate node and loop analysis: (*a*) typical node isolated for study; (*b*) typical loop isolated for study.

source. When this is done, the right side of the preceding equation will contain the dependent variables (voltages) in addition to independent current sources. These voltage terms can then be transposed to the left side of the matrix equation. The result is the addition of terms in the circuit matrix that make the matrix nonsymmetric.

Solution of Nodal Equations. In dc and ac (sinusoidal steady-state) circuits, the Y terms are numerical terms. Calculators that handle matrices and mathematical software programs for computers permit rapid solutions. Ordinary determinant methods also suffice. The result will be a set of values for the various voltages, all determined with respect to the reference node voltage. If the terms in the equation are generalized admittances (see Sec. 2.1.20 on Laplace transform analysis), then the solution will be a quotient of polynomials in the Laplace transform variable s. More is said about such solutions in that section.

Loop Current Analysis. Define a *loop* as a closed path in a circuit and a *loop current* as a current that flows around this path. See Fig. 2-32*b*, which shows one loop that has been isolated for attention, the associated loop current, and loop currents that flow in neighboring loops. It is noted that the current through any given element is found to be the difference between two loop currents if the circuit is planar, that is, can be drawn on a flat surface without crossing wires. (If the circuit is nonplanar, the technique is still valid, but it can become more complex, and some element currents will be composed of three or more loop currents.) The elements labeled Z are called impedances.

Kirchhoff's voltage law written around the loop states that

$$v_{k-2} + v_{k-1} + v_k + v_{k+1} + v_{k+2} = V_{in,k} \tag{2-80}$$

An expression for each of the voltages can be written

$$v_{k+1} = (i_k - i_{k+1})Z_{k+1} \tag{2-81}$$

When all the equations that can be witten are written, collected, and organized into matrix format, the general result is

$$\begin{bmatrix} Z_{11} + Z_{12} + Z_{13} + \cdots & -Z_{12} & -Z_{13} & \cdots \\ -Z_{21} & Z_{21} + Z_{22} + Z_{23} + \cdots & -Z_{23} & \cdots \\ -Z_{31} & -Z_{32} & Z_{31} + Z_{32} + Z_{33} + \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \end{bmatrix} \tag{2-82}$$

where the square matrix describes the circuit completely, the column matrix (vector) of currents describes the dependent variables which are the loop currents, and the column matrix of voltages describes the forcing function voltages that act in each loop.

Loop Current Analysis with Controlled Sources. If a controlled source is present, it is most convenient to use the Thevenin-Norton theorem to convert the controlled source to a current-controlled voltage source. When this is done, the right side of the preceding equation will contain the dependent variables (currents) in addition to independent voltage sources. These current terms can then be transposed to the left side of the matrix equation. The result is the addition of terms in the circuit matrix that make the matrix nonsymmetric.

Solution of Loop Current Equations. In D.C. and A.C. (sinusoidal steady-state) circuits, the Z terms are numerical terms. Calculators that handle matrices and mathematical software programs for computers facilitate the numerical work. Ordinary determinant methods also suffice. The result will be a set of values for the various loop currents, from which the actual element currents can be readily obtained. If the terms in the equation are generalized admittances (see Sec. 2.1.20 on Laplace transform analysis), then the solution will be a quotient of polynomials in the Laplace transform variable s. More is said about such solutions in those paragraphs.

Sinusoidal Steady-State Example. Figure 2-33 shows a circuit with a current source, two resistors, two capacitors, and one inductor. (The network is scaled.) The current source has a frequency of 2 rad/s and is sinusoidal. Figure 2-33*b* shows the circuit prepared for phasor analysis. The equations that follow show the writing of KCL equations for two voltages and their solution, which is shown as a phasor and as a time function.

$$2 = V_1(1 + j2.00 - j0.25) - V_2(-j0.25) \tag{2-83}$$

$$0 = -V_1(-j0.25) + V_2(1 + j2.00 - j0.25) \tag{2-84}$$

$$V_2 = 0.0615 - j0.1077 = 0.124e^{j(2.6224)} \text{ (angle in radians)} \tag{2-85}$$

$$V_2(t) = 0.1240\cos(2t + 2.6224) = 0.1240\cos(2t + 150.25°) \tag{2-86}$$

Computer Methods. The rapid development of computers in the last few years has led to the development of many programs written for the purpose of analyzing electric circuits. Because of their rapid analysis capability, they also are effective in design of new circuits. Programs exist for personal

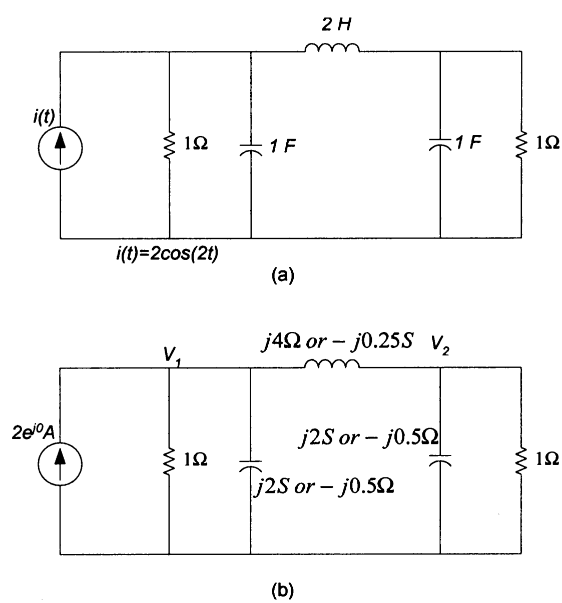


FIGURE 2-33 Sinusoidal steady-state analysis example: (*a*) circuit with sinusoidal source and two nodes; (*b*) phasor domain equivalent circuit.

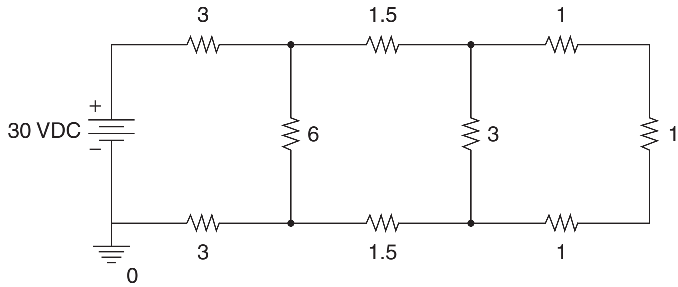


FIGURE 2-34 PSPICE circuit.

computers, minicomputers, and mainframe computers. Probably the most popular is SPICE, which is an acronym for Simulation Program with Integrated Circuit Emphasis. The personal computer version of this program is PSPICE. Most of these programs are in the public domain in the United States. It is convenient to discuss how a circuit is described to a computer program and what data are available in an analysis. Figures 2-34 and 2-35 show a sample PSPICE circuit.

SPICE Circuit Description. The analysis of a circuit with SPICE or another program requires the analyst to describe the circuit completely. Every node is identified, and each branch is described by type, numerical value, and nodes to which it is connected $e^{j(\alpha_1+\alpha_2)}$. Active devices such as transistors and operational amplifiers can be included in the description, and the program library contains complete data for many commonly used electronic elements.

SPICE Analysis Results. The analyst has a lot of control over what analysis results are computed. If a circuit is resistive, then a D.C. analysis is readily performed. This analysis is easily expanded to do a sensitivity analysis, which is a consideration of how results change when certain components change. Further, such analyses can be done both for linear and nonlinear circuits.

If the analyst wishes, a sinusoidal steady-state analysis is then possible. This includes *small-signal analysis,* a consideration of how well circuits such as amplifiers amplify signals which appear as currents or voltages. A frequency response is possible, and the results may be graphed with a variety of independent variables.

Other possible analyses include noise analyses—a study of the effect of electrical noise on circuit performance—and distortion analyses. Still others include transient response studies, which are most important in circuit design. The results may be graphed in a variety of useful ways. References give useful information.

Numerical example for the small-signal analysis is shown in Fig. 2-36.

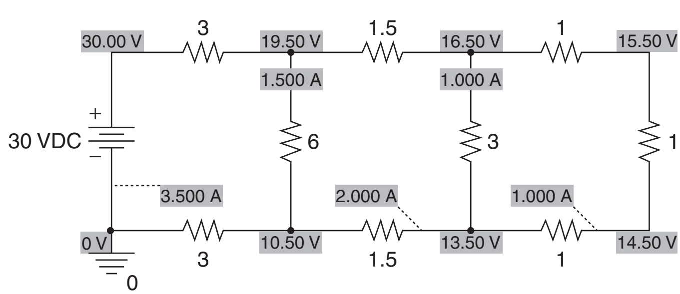


FIGURE 2-35 PSPICE analysis.

Additional Programs. A major advantage of these computer methods is that they work well for all types of circuit—low or high power, low or high frequency, power or communications. This is true even though the program's name might indicate otherwise. However, in some types of analysis, special programs have been developed to facilitate design and analysis. For example, power systems are often described by circuits that have more than 1000 nodes but very few nonzero entries in the circuit matrix. These special characteristics have led to the development of efficient programs for such studies. Several references address these issues.

$$V_{source} := 30\ \text{V}$$

$$R_1 := 3\ \Omega \qquad R_4 := 1.5\ \Omega \qquad R_7 := 1\ \Omega$$

$$R_2 := 6\ \Omega \qquad R_5 := 3\ \Omega \qquad R_8 := 1\ \Omega$$

$$R_3 := 3\ \Omega \qquad R_6 := 1.5\ \Omega \qquad R_9 := 1\ \Omega$$

$$R_{loop\ 3} := R_7 + R_8 + R_9$$

$$R_{loop\ 2} := R_4 + R_6 + \frac{1}{\frac{1}{R_5} + \frac{1}{R_{loop3}}}$$

$$R_{loop\ 1} := R_1 + R_3 + \frac{1}{\frac{1}{R_2} + \frac{1}{R_{loop2}}}$$

$$I := \frac{V_{source}}{R_{loop\ 1}}$$

$$I = 3.5\ \text{A}$$

FIGURE 2-36 MathCAD equation.

2.1.16 Electric Energy Distribution in 3-Phase Systems

General Note. In most of the world, large amounts of electric energy are distributed in 3-phase systems. The reasons for this decision include the fact that such systems are more efficient than single-phase systems. In other words, they have reduced losses and use materials more efficiently. Further, it can be shown that a 3-phase system distributes electric power at a constant rate, not at the time-varying rate shown earlier for single-phase systems. It is convenient to consider balanced systems and unbalanced systems separately. Also, both loads and sources need to be considered.

Balanced 3-Phase Sources. A 3-phase source consists of three voltage sources that are sinusoidal, equal in magnitude, and differ in phase by 120°. Thus, the set of voltages shown below is a balanced 3-phase source, shown both in time and phasor format.

$$
\begin{aligned}
v_{ab} &= V_m \cos(377t) && V_m e^{j0} \\
v_{bc} &= V_m \cos(377t - 120°) && V_m e^{-j120°} \qquad (2\text{-}87) \\
v_{ca} &= V_m \cos(377t + 120°) && V_m e^{+j120°}
\end{aligned}
$$

(In these expressions, peak values have been used rather than effective values. Further, degrees and radians are mixed, which is commonly done for the sake of clarity and convention but which can lead to numerical errors in calculators and computers if not reconciled.) Note that the sum of the three voltages is zero.

These three sources may be connected in either of the two ways to form a balanced system. One is the wye (star or tee) connection and the other is the delta (mesh or pi) connection. Both are shown in Fig. 2-37. It is noted that in the wye connection, a fourth point is needed, which is labeled O. The terminals labeled

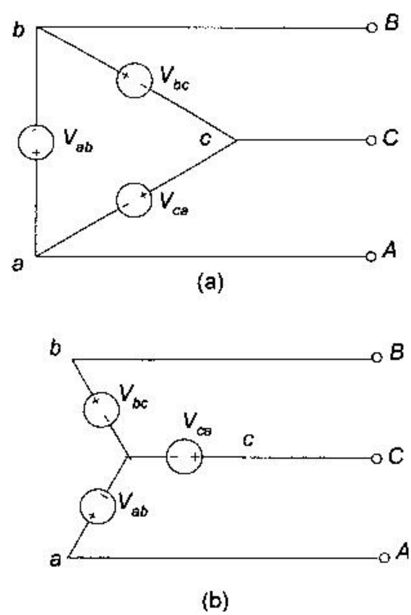


FIGURE 2-37 3-Phase source connections: *(a)* delta connection; *(b)* wye connection.

A, *B*, and *C* are called the *lines* (as opposed to *phases*). In the delta system, it is readily shown that, in phasor notation,

$$V_{AB} = V_{ab} = V_m^{e^{j0}}$$
$$V_{AB} = V_{bc} = V_m^{e^{j-120^\circ}} \qquad (2\text{-}88)$$
$$V_{CA} = V_{ca} = V_m^{e^{j+120^\circ}}$$

while in the wye system,

$$V_{AB} = V_{ab} - V_{bc} = V_{ao} - V_{bo} = \sqrt{3}V_m e^{j30^\circ}$$
$$V_{BC} = V_{bc} - V_{ca} = V_{bo} - V_{co} = \sqrt{3}V_m e^{j-90^\circ} \qquad (2\text{-}89)$$
$$V_{AB} = V_{ca} - V_{ab} = V_{co} - V_{ao} = \sqrt{3}V_m e^{j+150^\circ}$$

Thus, it is seen that in a delta system the line voltages are equal to the phase voltages. In a wye system, the line voltages are increased in magnitude by $\sqrt{3}$ and are shifted in phase by 30°. In a similar fashion, it is readily shown that the line currents in a wye system are equal to the phase currents, while in a delta, the line currents are increased by $\sqrt{3}$ and are shifted in phase by 30°.

Balanced Loads. A balanced 3-phase load is a set of three equal impedances connected either in wye or delta. Equations (2-74) and (2-75) may be used to convert from one to the other if needed.

Power Delivery, Balanced System. The power delivered from a balanced 3-phase source to a balanced 3-phase load is given by

$$P_{av} = \sqrt{3}V_{\text{line}}I_{\text{line}}\cos\phi$$

where $\cos\phi$ is, as before, the power factor of the load.

Unbalanced System. A 3-phase system that has either a nonsymmetric load or sources that differ in magnitude or whose phase difference is other than 120° is said to be unbalanced. Such circuits may be analyzed by any conventional method for circuit analysis.

Power Measurement in 3-Phase Systems. The power delivered to a 3-phase load, whether the system is balanced or unbalanced, may be measured with two wattmeters connected as shown in Fig. 2-38.

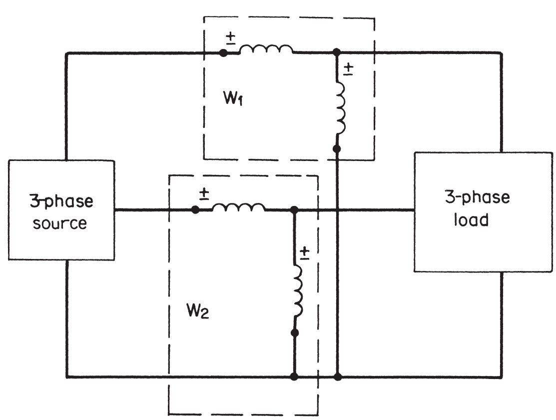


FIGURE 2-38 3-Phase system.

The total power is the sum of the readings of the two meters. If the power factors are small, one meter reading may be negative.

2.1.17 Symmetric Components

Resolution of an Unbalanced 3-Phase System into Balanced Systems. Let the three cube roots of unity, 1, $e^{j(2x/3)}$, $e^{j(4x/3)}$, be 1, a, a^2, where $j = \sqrt{-1}$,

$$a = 1 < 120° = -0.5 + j0.866$$

and

$$a = 1 < 120° = -0.5 - j0.866$$

Any three vectors, Q_a, Q_b, Q_c (which may be unsymmetric or unbalanced, that is, with unequal magnitudes or with phase differences not equal to 120°) can be resolved into a system of three equal vectors, Q_{a_0}, Q_{a_1}, Q_{a_2} and two symmetrical (balanced) 3-phase systems Q_{a_0}, $a^2Q_{a_1}$, Q_{a_2} and Q_{a_0}, aQ_{a_1}, Q_{a_2}, the first of which is of positive-phase sequence and the second of negative-phase sequence. Thus

$$\begin{aligned} Q_a &= Q_{a_0} + Q_{a_1} + Q_{a_2} \\ Q_b &= Q_{a_0} + a^2Q_{a_1} + aQ_{a_2} \\ Q_c &= Q_{a_0} + aQ_{a_1} + a^2Q_{a_2} \end{aligned} \tag{2-90}$$

The values of the component vectors are

$$\begin{aligned} Q_{a_0} &= 1/3(Q_a + Q_b + Q_c) \\ Q_b &= 1/3(Q_a + aQ_b + a^2Q_c) \\ Q_c &= 1/3(Q_a + a^2Q_b + aQ_c) \end{aligned} \tag{2-91}$$

The three equal vectors Q_{a_0} are sometimes called the *residual quantities* or the *zero-phase,* or *uniphase, sequence system*. Any of the vectors Q_a, Q_b, or Q_c may have the value zero. If two of them are zero, the single-phase system may be resolved into balanced 3-phase systems by the preceding equations. The symbol Q may denote any vector quantity such as voltage, current, or electric charge.

There are similar relations for n-phase systems. See "Method of Symmetrical Coordinates Applied to the Solution of Polyphase Networks," by C. L. Fortescue, *Trans. AIEE,* 1918, p. 1027.

Short-Circuit Currents. The calculation of short-circuit currents in 3-phase power networks is a common application of the method of symmetrical components. The location of a probable short circuit of fault having been selected, three networks are computed in detail from the neutrals to the fault, one for positive, one for negative, and one for zero-phase sequence currents. The three phases are assumed to be identical, in ohms and in mutual effects, except in the connection of the fault itself. Let Z_1, Z_2, and Z_0 be the ohms per phase between the neutrals and the fault in each of the networks, including the impedance of the generators.

Then, for a line-to-ground fault

$$I_{a_1} = I_{a_2} = I_{a_3} = \frac{I_c}{3} = \frac{V_a}{Z_1 + Z_2 + Z_0} \tag{2-92}$$

where V_a is the line-to-neutral voltage and I_{a_1} is the positive-phase-sequence current flowing to the fault in phase *a* and similarly for I_{a_2} and I_{a_0}. I_a is the total current flowing to the fault in phase *a*.

The component currents in phases *b* and *c* are derived from those in phase *a*, by means of the relations $I_{b_1} = a^2 I_{a_1}$, $I_{b_2} = aI_{a_2}$, $I_{b_0} = I_{a_0}$, $I_{c_1} = aI_{a_1}$, $I_{c_2} = a^2 I_{a_2}$, $I_{b_0} = I_{a_0}$. Each of the component currents divides in the branches of its own network according to the impedance of that network. Thus, each of the component currents, and therefore, the total current, at any part of the power system can be determined.

For a line-to-line fault between phases *b* and *c*

$$I_{a_1} = -I_{a_2} = \frac{V_a}{Z_1 + Z_2} \tag{2-93}$$

and

$$I_{a_0} = 0 \tag{2-94}$$

For a double line-to-ground fault between phases *b* and *c* and ground

$$I_{a_1} = \frac{V_a}{Z_1 + \dfrac{Z_2 Z_0}{Z_2 + Z_0}}$$

$$I_{a_0} = \frac{I_{a_1} Z_2}{Z_2 + Z_0}$$

and

$$I_{a_2} = -I_{a_1} - I_{a_0} \tag{2-95}$$

If there is no current in the power system before the fault occurs, the voltage V_a of every generator is the same in magnitude and phase. Such a condition often is assumed in calculated circuit-breaker duty and relay currents, although the effects of loads on the system can be included in the analysis.

In calculating power-system stability, however, it must be assumed that current exists in the lines before the fault occurs. The voltage V_a becomes the positive-sequence voltage at the point of fault before the fault occurs. A practical method of computing the positive-sequence current under fault conditions is to leave the positive-sequence network unchanged, with each generator at its own voltage and phase angle. The equivalent *Z* of the network need not be computed. Certain 3-phase impedances are inserted between line and neutral at the location of the fault. For a single line-to-line ground fault, $Z_2 + Z_0$ is inserted; for a line-to-line fault, Z_2 is inserted; and for a double line-to-ground fault, $Z_2 Z_0/(Z_2 + Z_0)$ is inserted. This gives one phase of an equivalent balanced 3-phase circuit for which the positive-sequence currents driven by all the generators in all the branches under fault conditions can be found by means of a network analyzer or computed on a digital computer. The power transmitted after the fault occurs can be determined from these positive-sequence currents.

If it is desired to find the negative-sequence and zero-sequence currents (some relays are operated by the latter), they can be computed from Eqs. (2-92) to (2-95) that do not involve *V*a, after finding I_{a_1} to the fault.

The impedance Z_f of each arc is mainly resistance. It may be brought into the computation. For single line-to-ground and double line-to-ground faults, Z_f is added to each of Z_1, Z_2, and Z_0. For line-to-line faults, Z_f is added to Z_2 only.

Load Studies. In calculations relating to the steady-state operation of power systems, in which it is desired to determine the voltage, power, reactive power, etc. at various points, the loads may be designated by kilowatts and kilovars rather than by impedances. The effect of the impedance of the transmission and distribution lines, transformers, etc. of the network can be computed. The modern method is a process of iterations using a digital computer for the calculations. The division of current in branches, the voltage at various points, and the required ratings of synchronous capacitors can be determined.

Conditions can be estimated at one or two points and a solution for the rest of the network can be calculated on the basis of these assumptions. If the assumptions are not correct, discrepancies will appear at the end of the work. For instance, two different voltages may be obtained for the same point, one calculated before and one after going around a loop of the network. The necessary correction to the first estimates may be based on the discrepancies, thus giving successive approximations which are improvements on the preceding ones.

2.1.18 Additional 3-Phase Topics

Voltage Drop in Unsymmetrical Circuits. The voltage drop due to resistance, self-inductance, and mutual inductance in any conductor of a group of long, parallel, round, nonmagnetic conductors forming a single-phase or polyphase circuit, and with one or more conductors connected electrically in parallel, may be calculated by summing the flux due to each conductor up to a certain large distance u. The vectorial sum of all the currents is zero in a complete system of currents in the steady state, and the quantity u cancels out, so the result is the same no matter how large u may be. The currents may be unbalanced, and in addition, the arrangement of the conductors may be unsymmetrical.

The voltage drop in any conductor a of a group of round conductors a, b, c, ... is

$$I_a R_a = j0.2794(I_a \log_{10} S_a + I_b \log_{10} S_{ab} + I_c \log_{10} S_{ac} + \cdots) \quad \text{V/mi at 60 Hz}$$

where $I_a + I_b + I_c = 0$, the values of the currents being complex quantities; R_a is resistance of conductor a, per mile; G_a is self-geometric mean distance of conductor a; G_a is axial spacing between conductors a and b, etc. The values of G and S should be in the same units.

Armature Windings. The armature winding of a 3-phase generator or motor is an important type of electric circuit. Windings consisting of diamond-shaped coils, with two coil sides per slot, are connected in groups of coils, three groups or phase belts being opposite each pole. In general, the number of slots per pole per phase is a fraction equal to the average number of coils per phase belt. There are a larger number and a smaller number of coils per phase belt, differing by 1. The winding is usually found to be divided into repeatable sections of several poles each, the sections being duplicates of each other.

The number of poles in a section is found by writing the fraction equal to the number of slots divided by the number of poles and canceling factors to the extent possible. The denominator is the number of poles per section, and the numerator is the number of slots per section.

If the final value of the numerator is not divisible by 3, a balanced 3-phase winding cannot be made, since the windings for phases a, b, and c in a section each require the same number of slots, and they must be duplicates except for the phase shift of 120°. This gives rise to the rule for balanced 3-phase windings that the factor 3 must occur at least one more time in the number of slots than in the number of poles.

It can be shown that the slots of a repeatable section have phase angles which, when suitably drawn, are all different and equidistant. They fill the space of 180 electrical degrees like the blades of a Japanese fan. The angle between the vectors in this fan is

$$\beta = \frac{180}{\text{slots per section}} \quad \text{deg} \tag{2-96}$$

The vectors lying from 0° to 59° may be assigned to phase a or $-a$, those from 60° to 119° to phase $-c$ or c, and those from 120° to 179° to phase b or $-b$.

The phase angles for the upper coil sides of the slots should be tabulated to indicate the proper connections of the winding. Since the diamond coils are all alike, the total resulting voltage developed in the lower coil sides of a phase is a duplicate of that developed in the upper coil sides and can be added on by means of the pitch factor.

The phase angle between two adjacent slots is

$$\frac{\text{Poles per section} \times 180}{\text{Slots per section}} = q\beta \quad \text{deg} \tag{2-97}$$

where q is the number of poles in the repeatable section. From this, the phase angle for every slot in the section can be written. To save numerical work, especially where β is a fractional number of degrees, the angles may be expressed in terms of the angle β, as given in the example below. They may be expressed in degrees and fractions of a degree, but decimal values of degrees should not be used in this part of the work. The required accuracy is obtained by using fractions instead of decimals. Appropriate multiples of 180° should be subtracted to keep the angles less than 180°, thus indicating the relative position of each coil side with respect to the nearest pole. When an odd number times 180° has been subtracted, the coil side is tabulated as $-a$ instead of a, etc., since it will be opposite a south pole when a is opposite a north pole. The terminals of a coil marked $-a$ are reversed with respect to the terminals of a coil marked a with which it is in series.

Example 21 slots per repeatable section; 5 poles per section; 11/5 slots per pole per phase.

$$\beta = \frac{180}{21} = 8\frac{4}{7} \qquad \text{deg} \qquad \text{[by Eq. 2-96]}$$

It is more convenient in this case to express the angles in terms of β rather than by fractions of degrees. Note that $21\beta = 180°$ and $7\beta = 60°$. The range for coils to be marked $\pm a$ is from 0 to 6β, inclusive; coils marked $\pm c$ from 7β to 13β; and coils marked $\pm c$ from 14β to 20β. Subtract multiples of $21\beta = 180°$. The angle between two adjacent slots is $q\beta = 5\beta$.

Tabulation of Phase Angles

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	[22]
0	5β	10β	15β	20β	$(25\beta)4\beta$	9β	14β	19β	$(24\beta)3\beta$	8β	13β	18β	$(23\beta)2\beta$	7β	12β	17β	$(22\beta)\beta$	6β	11β	16β	$[(21\beta)0]$
a	a	−c	b	b	−a	c	−b	−b	a	−c	−c	b	−a	c	c	−b	a	a	−c	b	[−a]

The seven vectors of phase a make a regular fan covering $7\beta = 60°$.

The resulting terminal voltage produced by the coils of phase a is equal to the numerical sum of the voltages in those coils multiplied by the "distribution factor"

$$\frac{\sin(n\beta/2)}{n\sin(\beta/2)} \tag{2-98}$$

where n is the number of vectors in the regular fan covering 60° and β is the angle between adjacent vectors, given by Eq. (2-95). The number n is large, and the perimeter approaches the arc of a circle. Equation (2-97) is of the same form as the formula for breadth factor, which also is based on a vector diagram that is a regular fan.

The distribution factor for the winding of the foregoing example is

$$\frac{\sin \dfrac{7 \times 60°}{2 \times 7}}{7 \sin \dfrac{60°}{2 \times 7}} = \frac{\sin 30°}{7 \sin 4\frac{2°}{7}} = \frac{0.5}{7 \times 0.0746} = 0.956$$

Other possible balanced 3-phase windings for this example could be specified by having some of the vectors of phase a lie outside the 60° range. This would result in a lower distribution factor. The voltage, and hence the rating of the machine, would be lower by 2% or more than in the case described. The canceling of the harmonic voltages would apparently not be improved, and there would be no advantage to compensate for the reduction in kVA rating, which would correspond to a loss or waste of 2% or more of the cost of the machine.

2.1.19 Two Ports

Two Ports. A common use of electric circuits is to connect a source of electric energy or information to a load, often in such a way as to modify the signal in a prescribed way. In its basic form, such a circuit has one pair of input terminals and one pair of output terminals. Each of the four

FIGURE 2-39 2-Port model with voltage and current definitions.

wires will have a current flow in or out of the circuit. Between any pair of terminals there will be a voltage.

If the structure of the circuit is such that the current flow into one terminal is equal to the current flow out of a second terminal, then that terminal pair is called a *port*. The circuit of Fig. 2-39 has two such pairs and is called a *2-port*. The circuit variables can be completely characterized at the terminals by the two currents and two voltages indicated in Fig. 2-39.

The last part of this section, "Filters," is devoted to a special type of 2-port called a *filter*. In the section, a variety of relations among the terminal voltages and currents will be discussed. Six such sets are found to be useful. They are known as the open-circuit impedance parameters, short-circuit admittance parameters, hybrid parameters (two types), and transmission-line parameters (two types).

Open-Circuit Impedance Parameters. If the two currents are considered to be independent variables and the two voltages are dependent variables, then this pair of equations can be written as

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \quad [v] = [z][i] \tag{2-99}$$

The four numbers or functions in the square matrix characterize the network. They may be computed by any conventional method of circuit analysis and often are computed directly by computer-based software. They also may be measured. For example, the term z_{21} can be computed or measured as the ratio v_2/i_1 when i_2 is set to zero if being computed or made zero by an appropriate open circuit when measurements are being taken.

Short-Circuit Admittance Parameters. If the two voltages are dependent variables and the two currents independent, then these equations can be written as

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad [i] = [y][v] \tag{2-100}$$

Measurements and computations follow principles similar to those of open-circuit impedance parameters.

Hybrid Parameters. Voltages and currents may be mixed in their roles as independent and dependent variables in two ways, as indicated:

$$\begin{bmatrix} v_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ v_2 \end{bmatrix} \quad (h \text{ parameters}) \tag{2-101}$$

$$\begin{bmatrix} i_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ i_2 \end{bmatrix} \quad (g \text{ parameters}) \tag{2-102}$$

Transmission-Line Parameters. The voltage and current at the input port may be used as independent variables with the output quantities as dependent variables, or the roles may be reversed.

TABLE 2-1 2-Port Conversions

$$[z] = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} \frac{y_{22}}{|y|} & -\frac{y_{12}}{|y|} \\ -\frac{y_{21}}{|y|} & \frac{y_{11}}{|y|} \end{bmatrix} = \begin{bmatrix} \frac{|h|}{h_{22}} & \frac{h_{12}}{h_{22}} \\ -\frac{h_{21}}{h_{22}} & \frac{1}{h_{22}} \end{bmatrix} = \begin{bmatrix} \frac{1}{g_{11}} & -\frac{g_{12}}{g_{11}} \\ \frac{g_{21}}{g_{11}} & \frac{|g|}{g_{11}} \end{bmatrix} = \begin{bmatrix} \frac{A}{C} & \frac{|TL|}{C} \\ \frac{1}{C} & \frac{D}{\gamma} \end{bmatrix} = \begin{bmatrix} \frac{\delta}{\gamma} & \frac{1}{\gamma} \\ \frac{|TLI|}{\gamma} & \frac{\alpha}{\gamma} \end{bmatrix}$$

$$[y] = \begin{bmatrix} \frac{z_{22}}{|z|} & -\frac{z_{12}}{|z|} \\ -\frac{z_{21}}{|z|} & \frac{z_{11}}{|z|} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{h_{11}} & -\frac{h_{12}}{h_{11}} \\ \frac{h_{21}}{h_{11}} & \frac{|h|}{h_{11}} \end{bmatrix} = \begin{bmatrix} \frac{|g|}{g_{22}} & -\frac{g_{12}}{g_{22}} \\ -\frac{g_{21}}{g_{22}} & \frac{1}{g_{22}} \end{bmatrix} = \begin{bmatrix} \frac{D}{B} & \frac{|TL|}{B} \\ -\frac{1}{B} & \frac{A}{B} \end{bmatrix} = \begin{bmatrix} \frac{\alpha}{\beta} & -\frac{1}{\beta} \\ \frac{|TLI|}{\beta} & \frac{\delta}{\beta} \end{bmatrix}$$

$$[h] = \begin{bmatrix} \frac{|z|}{z_{22}} & \frac{z_{12}}{z_{22}} \\ -\frac{z_{21}}{z_{22}} & \frac{1}{z_{22}} \end{bmatrix} = \begin{bmatrix} \frac{1}{y_{11}} & -\frac{y_{12}}{y_{11}} \\ \frac{y_{21}}{y_{11}} & \frac{|y|}{y_{11}} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} \frac{g_{22}}{|g|} & -\frac{g_{12}}{|g|} \\ -\frac{g_{21}}{|g|} & \frac{g_{11}}{|g|} \end{bmatrix} = \begin{bmatrix} \frac{B}{D} & \frac{|TL|}{D} \\ -\frac{1}{D} & \frac{C}{D} \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha} & -\frac{1}{\alpha} \\ -\frac{|TLI|}{\alpha} & \frac{\gamma}{\alpha} \end{bmatrix}$$

$$[g] = \begin{bmatrix} \frac{1}{z_{11}} & -\frac{z_{12}}{z_{11}} \\ \frac{z_{21}}{z_{11}} & \frac{|z|}{z_{11}} \end{bmatrix} = \begin{bmatrix} \frac{|y|}{y_{22}} & \frac{y_{12}}{y_{22}} \\ -\frac{y_{21}}{y_{22}} & \frac{1}{y_{22}} \end{bmatrix} = \begin{bmatrix} \frac{h_{22}}{|h|} & -\frac{h_{12}}{|h|} \\ -\frac{h_{21}}{|h|} & \frac{h_{11}}{|h|} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \frac{C}{A} & \frac{|TL|}{A} \\ \frac{1}{A} & \frac{B}{A} \end{bmatrix} = \begin{bmatrix} \frac{\gamma}{\delta} & -\frac{1}{\delta} \\ \frac{|TLI|}{\delta} & \frac{\beta}{\delta} \end{bmatrix}$$

$$[TL] = \begin{bmatrix} \frac{z_{11}}{z_{21}} & \frac{|z|}{z_{21}} \\ \frac{1}{z_{21}} & \frac{z_{22}}{z_{21}} \end{bmatrix} = \begin{bmatrix} -\frac{y_{22}}{y_{21}} & -\frac{1}{y_{21}} \\ -\frac{|y|}{y_{21}} & -\frac{y_{11}}{y_{21}} \end{bmatrix} = \begin{bmatrix} -\frac{|h|}{h_{21}} & -\frac{h_{11}}{h_{21}} \\ -\frac{h_{22}}{h_{21}} & -\frac{1}{h_{21}} \end{bmatrix} = \begin{bmatrix} \frac{1}{g_{21}} & \frac{g_{22}}{g_{21}} \\ \frac{g_{11}}{g_{21}} & \frac{|g|}{g_{21}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{\delta}{|TLI|} & \frac{\beta}{|TLI|} \\ \frac{\gamma}{|TLI|} & \frac{\alpha}{|TLI|} \end{bmatrix}$$

$$[TLI] = \begin{bmatrix} \frac{z_{22}}{z_{12}} & \frac{|z|}{z_{12}} \\ \frac{1}{z_{12}} & \frac{z_{11}}{z_{12}} \end{bmatrix} = \begin{bmatrix} -\frac{y_{11}}{y_{12}} & -\frac{1}{y_{12}} \\ \frac{|y|}{y_{12}} & -\frac{y_{22}}{y_{12}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{h_{12}} & -\frac{h_{11}}{h_{12}} \\ \frac{h_{22}}{h_{12}} & -\frac{|h|}{h_{12}} \end{bmatrix} = \begin{bmatrix} -\frac{|g|}{g_{12}} & -\frac{g_{22}}{g_{12}} \\ -\frac{g_{11}}{g_{12}} & -\frac{1}{g_{12}} \end{bmatrix} = \begin{bmatrix} \frac{D}{|TL|} & \frac{B}{|TL|} \\ \frac{C}{|TL|} & \frac{A}{|TL|} \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

$$\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} v_2 \\ i_2 \end{bmatrix} \qquad (h \text{ parameters}) \tag{2-103}$$

$$\begin{bmatrix} v_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \qquad (g \text{ parameters}) \tag{2-104}$$

2-Port Parameter Conversions. Any set of 2-port parameters may be converted to any other set through the use of Table 2-1. In this framework,

$$|y| = \det[y] = \det \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = y_{11}y_{22} - y_{12}y_{21} \tag{2-105}$$

With comparable interpretations for the other sets.

Equivalent Circuits for Two Ports. Equivalent circuits may be derived for any set of 2-port parameters. The process is shown for the $[h]$ parameters, but the technique is quite similar for the other combinations. Figure 2-40*a* shows the equivalent circuit.

Two-Port Analysis. From the equivalent circuit for a 2-port, circuit analysis is possible. Figure 2-40*b* shows a 2-port, with hybrid parameters, with a voltage source and a load resistor. Application of Kirchhoff's laws shows that

$$\frac{V_{\text{load}}}{V_{\text{source}}} = -\frac{R_L}{(R_{\text{source}} + h_{11})(1 + h_{22}R_{\text{load}}) + h_{12}R_{\text{load}}} \tag{2-106}$$

Similar analyses can be performed for any configuration that may arise.

Transmission Lines. The two sets of parameters denoted [*TL*] and [*TLI*] are called *transmission-line parameters* and are frequently computed or measured for power and communication transmission lines. Analysis with these parameters is substantially the same as that in the section on short-circuit admittance parameters above, though if the length of the line is more than about 10% of the wavelength involved, it is more convenient to use parameters based on standing-wave theory.

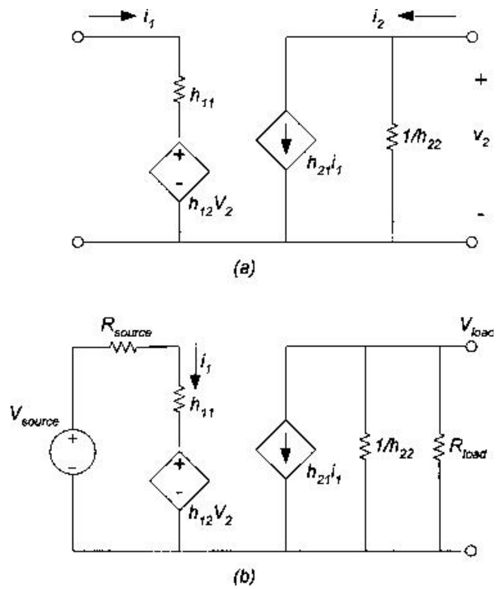


FIGURE 2-40 Transmission lines.

2.1.20 Transient Analysis and Laplace Transforms

Most transient analysis today is done with Laplace transform techniques, which provide the analyst a powerful method for finding both steady-state and transient computations simultaneously. It is necessary, however, to know initial conditions, the energy stored in capacitors and inductors.

Definition of a Laplace Transform. If a function of time $f(t)$ is known and defined for $t \geq 0$, then the (single-sided) Laplace transform is given by

$$Lf(t) = F(s) = \int_{0^-}^{\infty} f(t)e^{-st}\,dt \tag{2-107}$$

where s is a complex variable, $s = \sigma + j\omega$, chosen so that the integral will converge. In turn, σ is the real part of the variable, and ω is the imaginary part, but it becomes the frequency of sinusoidal functions measured in rad/s.

Laplace Transform Theorems. If the function $f(t)$ has the Laplace transform $F(s)$, then the theorems of Table 2-2 apply. In these theorems, the term $f(0^-)$ represents the initial condition, or the value of f at $t = 0$.

Laplace Transform Pairs. Table 2-3 presents a listing of the most common time functions and their Laplace transforms. These are sufficient for much of the analysis that is necessary. References include more extensive tables.

Initial Conditions. If a circuit has initial energy stored in it, that is, if any of the capacitors is charged or if any of the inductors has a nonzero current, then these conditions must be determined before a complete analysis can be done. In general, this will require knowledge of the circuit just before the circuit to be analyzed is connected. Normal circuit analysis methods can be used to determine the initial conditions.

TABLE 2-2 Laplace Transform Theorems

Operation	Theorem
Derivative	$\mathcal{L}\dfrac{df(t)}{dt} = sF(s) - f(0^-)$
nth-order derivative	$\mathcal{L}\dfrac{d^nf}{dt^n} = s^nF(s) - s^{n-1}f(0^-) - \cdots - f^{(n-1)}(0^-)$
Integral	$\mathcal{L}\displaystyle\int_{0^-}^{t} f(x)dx = \dfrac{F(s)}{s}$
Time shift	$\mathcal{L}[f(t - t_0)\, u\,(t - t_0)] = \epsilon_0^{-t} F(s)$
	where u is the unit step function
Frequency shift	$\mathcal{L}\epsilon^{-at}f(t) = F(s + a)$
Frequency scaling	$\mathcal{L}f(at) = \dfrac{1}{a}F\left(\dfrac{s}{a}\right), a > 0$
Initial value	$\lim\limits_{t\to 0} f(t) = \lim\limits_{t\to 0} sF(s)$
	provided the limit exists
	$\lim\limits_{t\to 0} f(t) = \lim\limits_{t\to 0} sF(s)$
	provided the limit exists
Constant multiplier	$\mathcal{L}kf(t) = k\mathcal{L}F(s)$
Addition	$\mathcal{L}[a_1f_1(t) + a_2f_2(t)] = a_1F_1(s) + a_2F_2(s)$
	a_1, a_2 are constants

Transfer Functions. The ratio of the Laplace transform of a response function to the Laplace transform of an excitation function, when initial conditions are zero, is called a *transfer function*. Any or all of the elements of the 2-port parameter matrices can be a transfer function in addition to being numerics. If the substitution $s = j\omega$ is made in a transfer function, then the new function is a function of (sinusoidal) signals of varying frequency. The frequency is measured in rad/s.

Example of Laplace Analysis. Figure 2-41 shows a 2-node circuit with an initial charge on one of the capacitors and the Laplace domain equivalent circuit. Equations (2-108) to (2-112) show Kirchhoff's current law equations for the circuit, a Laplace domain solution for one of the voltages, a partial fraction expansion of the solution, and finally, the inverse transform. (To reduce arithmetic complexity, the network is scaled.)

$$\frac{2s}{s^2 + 4} = V_1(s)\left(1 + \frac{1}{2s} + s\right) - V_2(s)\left(\frac{1}{2s}\right) \tag{2-108}$$

$$0 = -V_2(s)\left(\frac{1}{2s}\right) + V_1(s)\left(1 + \frac{1}{2s} + s\right) - 4 \tag{2-109}$$

$$V_2(s) = \frac{4s^4 + 4s^3 + 18s^2 + 17s + 8}{(s^2 + 4)(s + 1)(s^2 + s + 1)} \tag{2-110}$$

TABLE 2-3 Laplace Transform Pairs

Name	Time function f(t), t > 0⁻	Laplace transform $F(s)$
Unit impulse	$\delta(t)$	1
Unit step	$u(t)$	$\frac{1}{s}$
Unit ramp	t	$\frac{1}{s^2}$
nth-order ramp	t^n	$\frac{n!}{s^{n+1}}$
Exponential	ϵ^{-at}	$\frac{1}{s+a}$
Damped ramp	$t\epsilon^{-at}$	$\frac{1}{(s+a)^2}$
Cosine	$\cos \omega t$	$\frac{s}{s^2+\omega^2}$
Sine	$\sin \omega t$	$\frac{\omega}{s^2+\omega^2}$
Damped cosine	$\epsilon^{-at} \cos \omega t$	$\frac{s+a}{(s+a)^2+\omega^2}$
Damped sine	$\epsilon^{-at} \sin \omega t$	$\frac{\omega}{(s+a)^2+\omega^2}$

$$V_2(s) = \frac{1.8}{s+1} + \frac{-0.1077s - 0.1231}{s^2+4} + \frac{2.3077s + 0.23077}{s^2+s+1} \tag{2-111}$$

$$V_2(t) = 1.8e^{-1} - 0.1240 \cos(2t - 29.74°) + 2.5420e^{-t/2} \cos(0.8660t + 24.79°) \tag{2-112}$$

2.1.21 Fourier Analysis

Definition of Fourier Series. A periodic function $f(t)$ is defined as one that has the property

$$f(t) = f(t \pm nT) \tag{2-113}$$

where n is an integer and T is the period. If $f(t)$ satisfies the Dirichlet conditions, that is,

$f(t)$ has a finite number of finite discontinuities in the period T,

$f(t)$ has a finite number of maxima and minima in the period T,

the integral $\int_{t_0}^{t_0+T} |f(t)|\,dt$ exists. Then $f(t)$ can be written as a series of sinusoidal terms. Specifically,

$$f(t) = a_0 + \sum_{k=1}^{\infty} \left[a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t)\right] \tag{2-114}$$

The coefficients may be found with these equations:

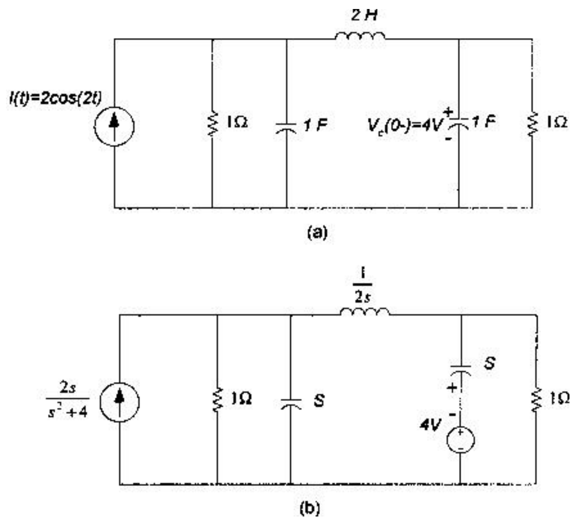


FIGURE 2-41 Example of Laplace transform circuit analysis: (*a*) 2-node circuit with one nonzero initial condition; (*b*) Laplace domain equivalent circuit equations and solution shown as Eqs. (2-108) to (2-112).

$$a_0 = \frac{1}{T}\int_{t_0}^{t_0+T} f(t)dt$$
$$a_k = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\cos(k\omega_0 t)dt \qquad (2\text{-}115)$$
$$b_k = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\sin(k\omega_0 t)dt$$

Evaluation of Fourier Coefficients. If an analytic expression for $f(t)$ is known, then the integrals can be used to evaluate the coefficients, which are then a function of the integral variable k. If the function $f(t)$ is known numerically or graphically, then numerical integration is required. Such integration is readily done with suitable computer software.

Effect of Symmetry. If the function $f(t)$ is even, that is, $f(t) = f(-t)$, then the expressions for the coefficients become

$$a_0 = \frac{2}{T}\int_0^{T/2} f(t)dt$$
$$a_k = \frac{4}{T}\int_0^{T/2} f(t)\cos(k\omega_0 t)\,dt \qquad (2\text{-}116)$$
$$b_k = 0 \qquad \text{for all } k$$

If $f(t)$ is odd, that is $f(t) = -f(-t)$, then

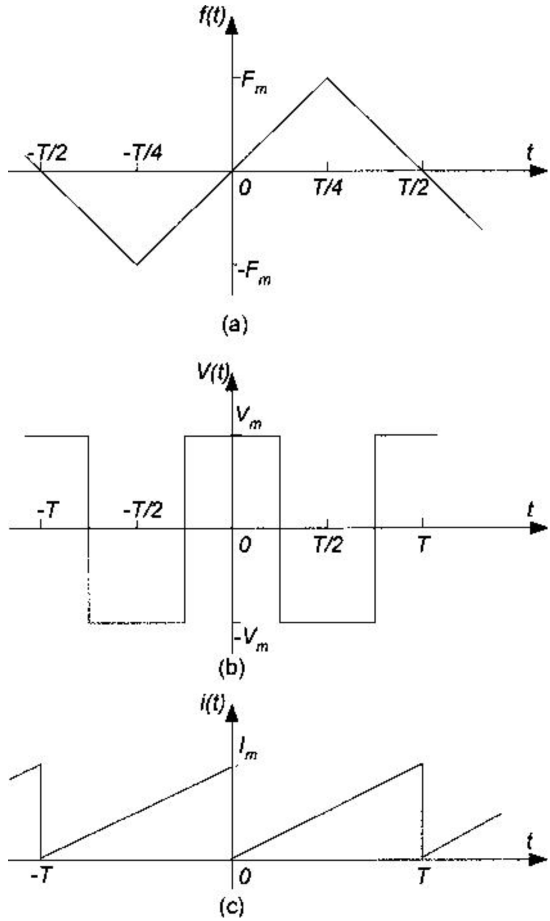


FIGURE 2-42 Waveforms for Fourier analyses [see Eqs. (2-118) to (2-120)]: (*a*) triangular wave, odd symmetry; (*b*) square wave, even symmetry; (*c*) ramp function.

$$a_k = 0 \qquad \text{for} \qquad k = 0, 1, 2, 3, \ldots$$

$$b_k = \frac{4}{T}\int_0^{T/2} \sin(k\omega_0 t)dt \tag{2-117}$$

Fourier Series Examples. Figure 2-42 shows three waveforms: a triangular signal written as an odd function, a square wave written as an even function, and a ramp function. For these three signals, the Fourier series are

$$f(t) = \frac{8F_m}{\pi^2}\sum_{n=1,3,5,\ldots}^{\infty}\frac{1}{n^2}\sin\left(\frac{n\pi}{2}\right)\sin(n\omega_0 t)$$

$$v(t) = \frac{4v_m}{\pi}\sum_{n=1,3,5,\ldots}^{\infty}\frac{1}{n}\sin\left(\frac{n\pi}{2}\right)\sin(n\omega_0 t) \tag{2-118}$$

$$i(t) = \frac{I_m}{2} - \frac{I_m}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin(n\omega_0 t)$$

Average Power Calculations. If the Fourier series given in Eq. (2-114) represents a current or voltage, then the effective value of this current or voltage is given by

$$F_{\text{eff}} = \sqrt{a_0^2 + \frac{1}{2}\sum_{1}^{\infty}(a_k^2 + b_k^2)} \qquad (2\text{-}119)$$

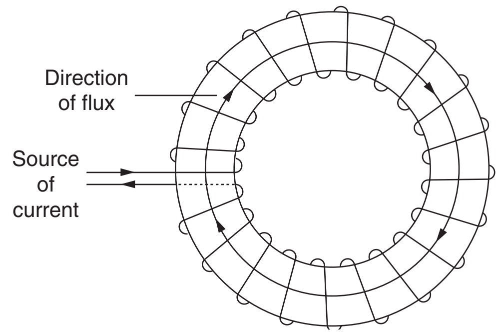


FIGURE 2-43 Closed magnetic circuit.

2.1.22 The Magnetic Circuit

The Simple Magnetic Circuit. A simple magnetic circuit is a uniformly wound torus ring (Fig. 2-43). The relation between the mmf F and the flux ϕ is similar to Ohm's law, namely,

$$F = R\phi \qquad \text{At} \qquad (2\text{-}120)$$

where R is called the *reluctance* of the magnetic circuit. The relation is sometimes written in the form

$$\phi = PF$$

where $p = 1/R$ is called the *permeance* of the magnetic circuit. Reluctance is analogous to resistance, and permeance is analogous to conductance of an electric circuit.

$$F = NI \qquad \text{At} \qquad (2\text{-}121)$$

where N is the number of turns of conductor around the magnetic circuit, as in Fig. 2-43, and I is the current in the conductor, in amperes.

Permeability and Reluctivity. The reluctance of a uniform magnetic path (Fig. 2-43) is proportional to its length I and inversely proportional to its cross section A.

$$R = v\frac{1}{A} \qquad \text{At/Wb} \qquad (2\text{-}122)$$

and

$$P = \mu\frac{A}{l} \qquad \text{Wb/At} \qquad (2\text{-}123)$$

In these expressions, v is called the *reluctivity* and μ the *permeability* of the material of the magnetic path, it being assumed that there is no residual magnetism. The dimensions l and A are in metric units. For a vacuum, air, or other nonmagnetic substance, the reluctivity and permeability are usually written v_0 and μ_0, and their values are $1/(4\pi \times 10^{-7})$ and $4\pi \times 10^{-7}$, respectively.

Magnetic Field Intensity. Magnetic field intensity H is defined as the mmf per unit length of path of the magnetic flux. It is known also as the *magnetizing force* or the *magnetic potential gradient*. In a uniform field,

$$H = \frac{F}{l} \qquad \text{At/m} \qquad (2\text{-}124)$$

In a nonuniform magnetic circuit,

$$H = \frac{\partial F}{\partial l} \qquad (2\text{-}125)$$

Inversely, for a uniform field,

$$F = HI \tag{2-126}$$

and for a nonuniform field,

$$F = \int H dl \tag{2-127}$$

By Ampere's law, when this integral is taken around a complete magnetic circuit,

$$F = \oint H dl = I \tag{2-128}$$

where I is the total current, in amperes, surrounded by the magnetic circuit. The circle on the integral sign indicates integration around the complete circuit. In Eqs. (2-126) to (2-128), it is presumed that H is directed along the length of l; otherwise, the factor $\cos\theta$ must be added to the product Hdl, where θ is the angle between H and dl.

Flux Density. Flux density B is the magnetic flux per unit area, the area being perpendicular to the direction of the magnetic lines of force. In a uniform field,

$$B = \frac{\phi}{A} \qquad \text{T(or Wb/m}^2\text{)} \tag{2-129}$$

Reluctances and Permeances in Series and in Parallel. Reluctances and permeances are added like resistances and conductances, respectively. That is, *reluctances are added when in series, and permeances are added when in parallel*. If several permeances are given connected in series, they are converted into reluctances by taking the reciprocal of each. If reluctances are given in a parallel combination, they are similarly converted into permeances.

Magnetization Characteristic or Saturation Curve. The magnetic properties of steel or iron are represented by a saturation or magnetization curve (Fig. 2-44). Magnetic field intensities H in ampere-turns per meter are plotted as abscissas and the corresponding flux densities B in teslas (webers per square meter) as ordinates.

The practical use of a magnetization curve may be best illustrated by an example. Let it be required to find the number of exciting ampere-turns for magnetizing a steel ring so as to produce in it a flux of 1.68 mWb. Let the cross section of the ring be 10.03 by 0.04 m and the mean diameter 0.46 m. Let the quality of the material be represented by the curve in Fig. 2-44. The flux density is $1.68 \times 10^{-3}/(0.03 \times 0.04) = 1.4$ Wb/m^2. For this flux density, the corresponding abscissa from the curve is about 18 At/m. The total required number of ampere-turns is then $18 \times \pi \times 46 = 2600$.

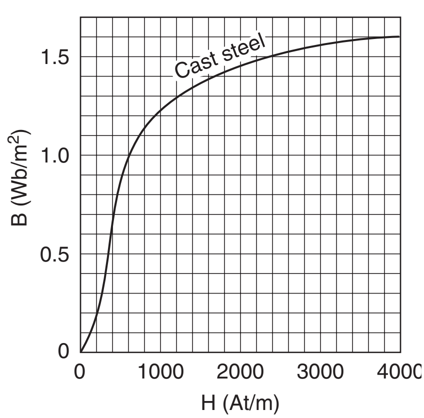


FIGURE 2-44 Typical *BH* curve.

For curves of various grades of steel and iron, see Sec. 4. The principal methods for experimentally obtaining magnetization curves will be found in Sec. 3.

Ampere-Turns for an Air Gap. In a magnetic circuit consisting of iron with one or more small air gaps in series with the iron, the magnetic flux density in each of the air gaps may be considered approximately uniform. If the length across a given air gap in the direction of the flux is l m, the ampere-turns required for that air gap is given by the equation

$$\text{At/m} = \frac{B(T)}{4\pi \times 10^{-7}} = 7.958 \times 10^5 B(T) \tag{2-130}$$

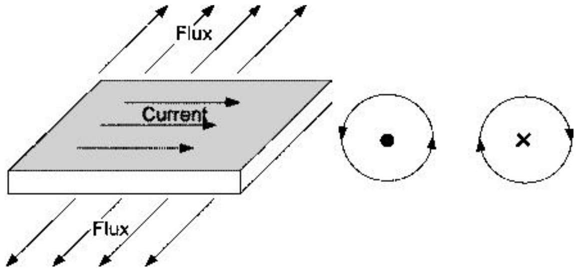


FIGURE 2-45 Relation between direction of current and flux.

The ampere-turns for each portion of iron, computed from iron magnetization curves such as Fig. 2-44, and the ampere-turns for the air gaps are added together to give the ampere-turns for the complete magnetic circuit.

Analysis of Magnetization Curve. Three parts are distinguished in a magnetization curve (Fig. 2-44): the lower, or nearly straight, part; die middle part, called the *knee* of the curve; and the upper part, which is nearly a straight line. As the magnetic intensity increases, the corresponding flux density increases more and more slowly, and the iron is said to approach saturation (see Sec. 4).

Magnetization per Unit Volume and Susceptibility. If a portion of ferromagnetic material is magnetized by an mmf, H At/m, the resulting flux density in teslas may be written as

$$B = \mu_0(H + M) \tag{2-131}$$

where M is the magnetization per unit volume of the material (see Sec. 4).

The ratio of M/H is symbolized by x and is called the *magnetic susceptibility*. It is the excess of the ratio of $B/\mu_0 H$ above unity, that is,

$$x = \frac{B}{\mu_0 H} - 1 \tag{2-132}$$

This is a dimensionless quantity. See Sec. 1.

The Right-Handed-Screw Rule. The direction of the flux produced by a given current is determined as shown in Fig. 2-45 (see also Fig. 2-43). If the current is established in the direction of rotation of a right-handed screw, the flux is in the direction of the progressive movement of the screw. If the current in a straight conductor is in the direction of the progressive motion of a right-handed screw, then the flux encircles this conductor in the direction in which the screw must be rotated in order to produce this motion. The dots in the figure indicate the direction of flux or current toward the reader, and the crosses away from the reader.

Fleming's Rules. The relative direction of flux, voltage, and motion in a revolving-armature generator may be determined with the right hand by placing the thumb, index, and middle fingers so as to form the three axes of a coordinate system and pointing the index finger in the direction of the flux (north to south) and the thumb in the direction of motion; the middle finger will give the direction of the generated voltage (Fig. 2-46). In the same way, in a revolving-armature motor, by using the left hand and pointing the index finger in the direction of the flux and the middle finger in the direction of the current in the armature conductor, the thumb will indicate the direction of

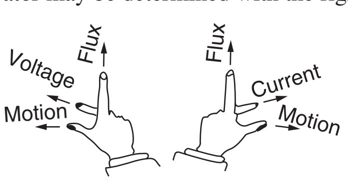


FIGURE 2-46 Flemming's generator and motor rules.

the force and, therefore, the resulting motion. These two rules, indicated in Fig. 2-46, are known as *Fleming's rules*.

Magnetic Tractive Force. The attracting force of a magnet is

$$F = \frac{1}{2}\frac{AB^2}{\mu_0} = \frac{AB^2}{8\pi \times 10^{-7}} \qquad \text{newtons} \tag{2-133}$$

where B is the flux density in the air gap, expressed in teslas (webers per square meter), and A is the total area of the contact between the armature and the core, in square meters. The mass that can be supported is dependent on the gravity field in which the mass and magnet are located.

Magnetic Force, or Torque. The mechanical force, or the torque, between two parts of a magnetic or electric circuit may in some cases be conveniently calculated by making use of the principle of *virtual displacements*. An infinitesimal displacement between the two parts is assumed. The energy supplied from the source of current is then equal to the mechanical energy for producing the motion, plus the change in the stored magnetic energy, plus the energy for resistance loss.

When the differential motion ds m of a part of a circuit carrying a current I A changes its self-inductance by a differential dL H, the mechanical force on that part of the circuit, in the direction of the motion, is

$$F = \frac{1}{2}I^2\frac{dL}{ds} \qquad \text{newtons} \tag{2-134}$$

When the motion of one coil or circuit carrying a current I_1 A changes its mutual inductance by a differential dM H with respect to another coil or circuit carrying a current I_2 A, the mechanical force on each coil or circuit, in the direction of the motion, is

$$F = I_1 I_2 \frac{dM}{ds} \qquad \text{newtons} \tag{2-135}$$

where ds represents the differential of distance in meters. For a discussion of self-inductance and mutual inductance L and M, see Sec. 2.1.6.

2.1.23 Hysteresis and Eddy Currents in Iron

The Hysteresis Loop. When a sample of iron or steel is subjected to an alternating magnetization, the relation between B and H is different for increasing and decreasing values of the magnetic intensity (Fig. 2-47). This phenomenon is due to irreversible processes which result in energy dissipation, producing heat. Each time the current wave completes a cycle, the magnetic flux wave also must complete a cycle, and the elementary magnets are turned. The curve *AefBcdA* in Fig. 2-47 is called the *hysteresis loop*.

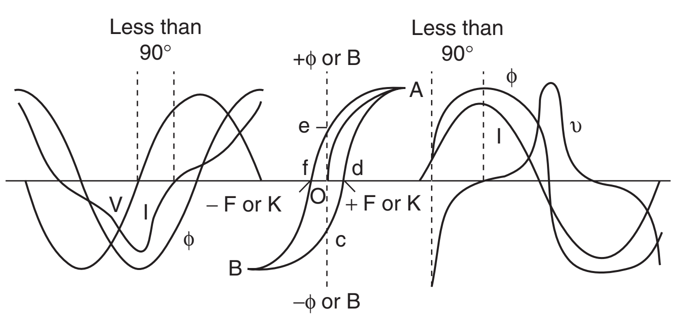


FIGURE 2-47 Periodic waves of current, flux, and voltage; hysteresis loop.

Retentivity. If the coil shown in Fig. 2-43 is excited with alternating current, the ampere-turns and consequently the mmf will, at any instant, be proportional to the instantaneous value of the exciting current. Plotting a *B-H* (or ϕ-*F*) curve (Fig. 2-47) for one cycle yields the closed loop *AefBcdA*. The first time the iron is magnetized, the *virgin*, or *neutral, curve OA* will be produced, but it cannot be produced in the reverse direction *AO* because when the mmf drops to zero there will always be some remaining magnetism ($+Oe$ or $-Oc$). This is called *residual magnetism*; to reduce this to zero, an mmf ($-Of$ or $+Od$) of opposite polarity must be applied. This mmf is called the *coercive force*.

Wave Distortion. In Fig. 2-47 the instantaneous values of the exciting current *I* (which is directly proportional to the mmf) and the corresponding values of the flux ϕ and voltage *V* (or *v*) are plotted against time as abscissas, beside the hysteresis loop. (a) If the voltage applied to the coil is *sinusoidal* (*V*, to the left), the current wave is distorted and displaced from the corresponding sinusoidal flux wave. The latter wave is in quadrature with the voltage wave. (b) If the *current* through the coil is *sinusoidal* (*I*, to the right), the flux is distorted into a flat-top wave and the induced voltage *v* is peaked.

Components of Exciting Current. The alternating current that flows in the exciting coil (Fig. 2-47) may be considered to consist of two components, one exciting magnetism in the iron and the other supplying the iron loss. For practical purposes, both components may be replaced by equivalent sine waves and phasors (Fig. 2-48) (see Sec. 2.1.11). We have

$$I_r = I\cos\theta = \text{power component of current}$$

$$P_h = IV\cos\theta = I_rV = \text{iron loss in watts} \tag{2-136}$$

$$I_m = I\sin\theta = \text{magnetizing current}$$

where *I* is the total exciting current, and *J* the angle of time-phase displacement between current and voltage.

Hysteretic Angle. Without iron loss, the current *I* would be in phase quadrature with *V*. For this reason, the angle $\alpha = 90 - \theta$ is called the *angle of hysteretic advance of phase*.

$$\sin\alpha = \frac{I_r}{I} = \frac{I_rN}{IV} = \frac{W\text{loss}}{VA} \tag{2-137}$$

In practice, the measured loss usually includes eddy currents, so the name *hysteretic* is somewhat of a misnomer.

The energy lost per cycle from hysteresis is proportional to the area of the hysteresis loop (Fig. 2-47). This is a consequence of the evaluation over a cycle of Eq. (2-13).

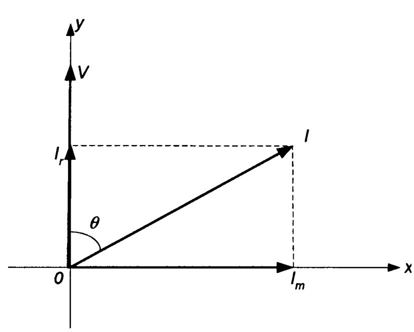


FIGURE 2-48 Components of exciting current; hysteretic angle.

Steinmetz's Formula. According to experiments by C.P. Steinmetz, the heat energy due to hysteresis released per cycle per unit volume of iron is approximately

$$W_h = \eta B_{\max}^{1.6} \tag{2-138}$$

The exponent of $B_{\max}$ varies between 1.4 and 1.8 but is generally taken as 1.6. Values of the hysteresis coefficient 11 are given in Sec. 4.

Eddy-Current Losses. Eddy-current losses are I^2R losses due to secondary currents (Foucault currents) established in those parts of the circuit which are interlinked with alternating or pulsating flux. Refer to

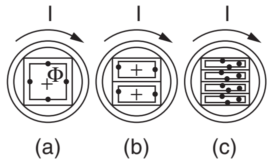


FIGURE 2-49 Section of a transformer core.

Fig. 2-49, which shows a cross section of a transformer core. The primary current I produces the alternating flux Φ, which by its change generates a voltage in the core; this voltage then sets up the secondary current i. Now, if the core is divided into two (b), four (c), or n parts, the voltage in each circuit is $v/2$, $v/4$, v/n, and the conductance is $g/2$, $g/4$, g/n, respectively. Thus, the loss per lamination will be $1/n^3$ times the loss in the solid core, and the total loss is $1/n^2$ times the loss in the solid core. When power is computed, the power is given by

$$P_h = \eta f \beta_{\max}^{1.6} \tag{2-139}$$

Eddy currents can be greatly reduced by laminating the circuit, that is, by making it up of thin sheets each electrically insulated from the others. The same purpose is accomplished by using separately insulated strands of conductors or bundles of wires.

A formula for the eddy loss in conductors of circular section, such as wire, is

$$P_e = \frac{(\pi r f B_{\max})^2}{4\rho} \qquad \text{W/m}^3 \tag{2-140}$$

where r is the radius of the wire in meters, f is the frequency in hertz (cycles per second), $B_{\max}$ is the maximum flux density in teslas, and p is the specific resistance in ohm meters.

A formula for the loss in sheets is

$$P_e = \frac{(\pi r f B_{\max})^2}{6\rho} \qquad \text{W/m}^3 \tag{2-141}$$

where t is the thickness in meters, f is the frequency in hertz (cycles per second), $B_{\max}$ is the maximum flux density in teslas, and ρ is the specific resistance in ohm meters. The specific resistance of various materials is given in Sec. 4.

Effective Resistance and Reactance. When an A.C. circuit has appreciable hysteresis, eddy currents, and skin effect, it can be replaced by a circuit of equivalent resistances and equivalent reactances in place of the actual ones. These effective quantities are so chosen that the energy relations are the same in the equivalent circuit as in the actual one. In a series circuit, let the true power lost in ohmic resistance, hysteresis, and eddy currents be P, and the reactive (wattless) volt-amperes, Q. Then the effective resistance and reactance are determined from the relations

$$i^2 r_{\text{eff}} = P \qquad i^2 x_{\text{eff}} = Q \tag{2-142}$$

In a parallel circuit, with a given voltage, the equivalent conductances and susceptance are calculated from the relations

$$e^2 g_{\text{eff}} = P \qquad e^2 b_{\text{eff}} = Q \tag{2-143}$$

Such equivalent electric quantities, which replace the core loss, are used in the analytic theory of transformers and induction motors.

Core Loss. In practical calculations of electrical machinery, the total core loss is of interest rather than the hysteresis and the eddy currents separately. For such computations, empirical curves are used, obtained from tests on various grades of steel and iron (see Sec. 4).

Separation of Hysteresis Losses from Eddy-Current Losses. For a given sample of laminations, the total core loss P, at a constant flux density and at variable frequency f, can be represented in the form

$$P = af + bf^2 \tag{2-144}$$

where af represents the hysteresis loss and bf^2 the eddy, or Foucault-current, loss, a and b being constants. The voltage waveform should be very close to a sine wave. If we write this equation for two known frequencies, two simultaneous equations are obtained from which a and b are determined.

It is convenient to divide the foregoing equation by f, because the form

$$\frac{P}{f} = a + bf \tag{2-145}$$

represents a straight line relating P/f and f. Known values of P/f are plotted against f as abscissas, and a straight line having the closest approximation to the points is drawn. The intersection of this line with the axis of ordinates gives a; b is calculated from the preceding equation. The separate losses are calculated at any desired frequency from af and bf^2, respectively.

2.1.24 Inductance Formulas

Inductance. The properties of self-inductance and mutual inductance are defined in Sec. 2.1.6. In these paragraphs, inductance relations for common geometries are given.

Torus Ring or Toroidal Coil of Rectangular Section with Nonmagnetic Core (Fig. 2-43). Inductance of a rectangular toroidal coil, uniformly wound with a single layer of fine wire, is

$$L = 2 \times 10^{-7} N^2 b \left(\ln \frac{r_2}{r_1}\right) \qquad \text{henrys} \tag{2-146}$$

where N equals the number of turns of wire on the coil, b is the axial length of the coil in meters, and r_2 and r_1 are the outer and inner radial distances in meters.

Torus Ring or Toroidal Coil of Circular Section with Nonmagnetic Core (Fig. 2-43). A toroidal coil of circular section, uniformly wound with a single layer of fine wire of N turns, has an inductance of

$$L = 4\pi \times 10^{-7} N^2 \left(g - \sqrt{g^2 - a^2}\right) \qquad \text{henrys} \tag{2-147}$$

where g is the mean radius of the toroidal ring and a is the radius of the circular cross section of the core, both measured in meters.

Inductance of a Very Long Solenoid. A solenoid uniformly wound in a single layer of fine wire possesses an inductance of

$$L = \frac{4\pi^2 \times 10^{-7} N^2 R^2}{S} \qquad \text{henrys} \tag{2-148}$$

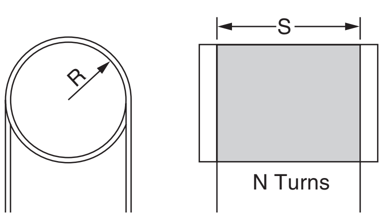


FIGURE 2-50 Cylindrical solenoid.

where R and S are the radius and length of the solenoid in meters, as illustrated in Fig. 2-50. The assumption is made that S is very large with respect to R.

Inductance of the Finite Solenoid. The inductance of a short solenoid is less than that given by Eq. (2-148), by a factor k (a dimensionless quantity). The inductance relation then is

$$L = k \frac{4\pi^2 \times 10^{-7} N^2 R^2}{S} \qquad \text{henrys} \tag{2-149}$$

where the values of k for various ratios of R and S are given in Fig. 2-51. Inductance relations for other configurations of coils are given by Boast (1964).

Inductance per Unit Length of a Coaxial Cable. For low-frequency applications, where skin effect is not predominant (uniform current density over nonmagnetic current-carrying cross sections), the inductance per unit length of a coaxial cable is

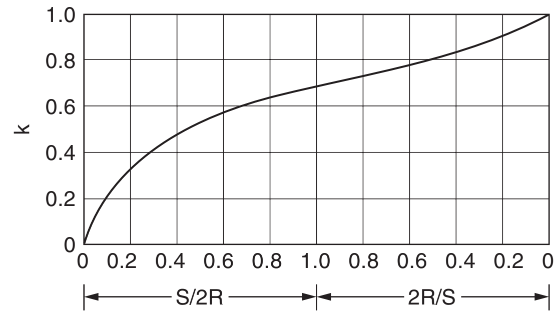


FIGURE 2-51 Factor k of Eq. (2-149). (*From W. B. Boast, Vector Fields, New York, Harper & Row, 1964.*)

$$L = l = \frac{10^{-7}}{2}[1 + 4\ln\frac{R^2}{R_1} + \frac{4R_3^4}{(R_3^2 - R_2^2)^2}\ln\frac{R_3}{R_2} - \frac{3R_3^2 - R_2^2}{R_3^2 - R_2^2}] \quad \text{H/m} \tag{2-150}$$

where R_1, R_2, and R_3 are the radii of the inner conductor, the inner radius of the outer conductor, and the outer radius of the outer conductor, in meters, respectively, as shown in Fig. 2-52. For very thin outer shells, the last two terms drop out of the equation, and for very small inner conductors, the first term becomes less important. For high-frequency applications, the first, third, and fourth terms are all suppressed, and for the extreme situation where all the current is essentially at the boundaries formed by R_1 and R_2, respectively, the inductance per unit length becomes

$$l = 2 \times 10^{-7}\ln(R_2/R_1) \quad \text{H/m} \tag{2-151}$$

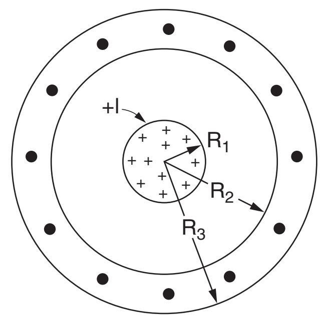


FIGURE 2-52 Coaxial cable.

Inductance of Two Long, Cylindrical Conductors, Parallel and External to Each Other. The inductance per unit length of two separate parallel conductors is

$$l = 10^{-7}\left(1 + 4\ln\frac{D}{\sqrt{R_1R_2}}\right) \quad \text{H/m} \tag{2-152}$$

where D is the distance between centers of the two cylinders and R_1 and R_2 are the radii of the conductor cross sections.

If $R_1 = R_2 = R$ and the skin-effect phenomenon applies as at very high frequencies, the inductance per unit length becomes

$$l = 4 \times 10^{-7}\left(1 + 4\ln\frac{D}{\sqrt{R_1R_2}}\right) \quad \text{H/m} \tag{2-153}$$

Inductance of Transmission Lines. The inductance relationships used in predicting the performance of power-transmission systems often involve the effects of stranded and bundled conductors operating in parallel, as well as configurations of these groups of current-carrying elements of one phase group of the system coordinated with similar groups constituting other phases, in polyphase systems. In such systems, the several current-carrying elements of a phase are considered mathematically as a cylindrical shell of current of radius D_s (meters) called the *self-geometric mean radius* of the phase, and the mutual distances (between the current in a particular phase and the other [return] currents in the other phases) are replaced by a distance D_m (meters) called the *mutual geometric mean distance* to the return. The inductance of all phases may be balanced by transposing the conductors over the length of the transmission line so that each phase occupies all positions equally in the length of the line.

The inductance per phase is then one-half as large as that of Eq. (2.153), that is,

$$l = 2 \times 10^{-7} 4 \ln(D_m/D_s) \qquad \text{H/m} \tag{2-154}$$

The references related to methods for computing the geometric mean distances D_m and D_s are available in Bibliography at the end of the section.

Leakage Inductance. In electrical apparatus, such as transformers, generators, and motors, in which the greater part of the flux is carried by an iron core, the difference between self-inductance and mutual inductance of the primary and secondary windings is small. This small difference is called *leakage inductance*. It is of great importance in the characteristics and operation of the apparatus and is usually calculated or measured separately. The loss in voltage in such apparatus, due to inductance, is associated with the leakage.

Magnetizing Current. The mutual inductance of the windings of apparatus with iron cores is not usually stated in henrys, but the effective alternating current required to produce the flux is stated in amperes and is called the *exciting current*. One component of this current supplies the energy corresponding to the core loss. The remaining component is called *magnetizing current.* Solenoids and other coils with only one winding are usually treated in a similar manner when they have iron cores. The exciting current usually does not have a sine-wave form. See Fig. 2-47.

2.1.25 Skin Effect

Real, or ohmic, resistance is the resistance offered by the conductor to the passage of electricity. Although the specific resistance is the same for either alternating or continuous current, the total resistance of a wire is greater for alternating than for continuous current. This is due to the fact that there are induced emfs in a conductor in which there is alternating flux. These emfs are greater at the center than at the circumference, so the potential difference tends to establish currents that oppose the current at the center and assist it at the circumference. The current is thus forced to the outside of the conductor, reducing the effective area of the conductor. This phenomenon is called *skin effect*.

Skin-Effect Resistance Ratio. The ratio of the A.C. resistance to the D.C. resistance is a function of the cross-sectional shape of the conductor and its magnetic and electrical properties as well as of the frequency. For cylindrical cross sections with presumed constant values of relative permeability μ_r and resistivity ρ, the function that determines the skin-effect ratio is

$$mr = \sqrt{\frac{8\pi^2 \times 10^{-7} f \mu_r r}{\rho}} \tag{2-155}$$

where r is the radius of the conductor and f is the frequency of the alternating current. The ratio of R, the A.C. resistance, to R_0, the D.C. resistance, is shown as a function of mr in Fig. 2.53.

Steel Wires and Cables. The skin effect of steel wires and cables cannot be calculated accurately by assuming a constant value of the permeability, which varies throughout a large range during every cycle. Therefore, curves of measured characteristics should be used. See *Electrical Transmission and Distribution Reference Book,* 4th ed., 1950.

Skin Effect of Tubular Conductors. Cables of large size are often made so as to be, in effect, round, tubular conductors. Their effective resistance due to skin effect may be taken from the curves of Sec. 4. The skin-effect ratio of square, tubular bus bars may be obtained from semiempirical formulas in the paper "A-C Resistance of Hollow, Square Conductors," by A. H. M. Arnold, *J. IEE (London),* 1938,

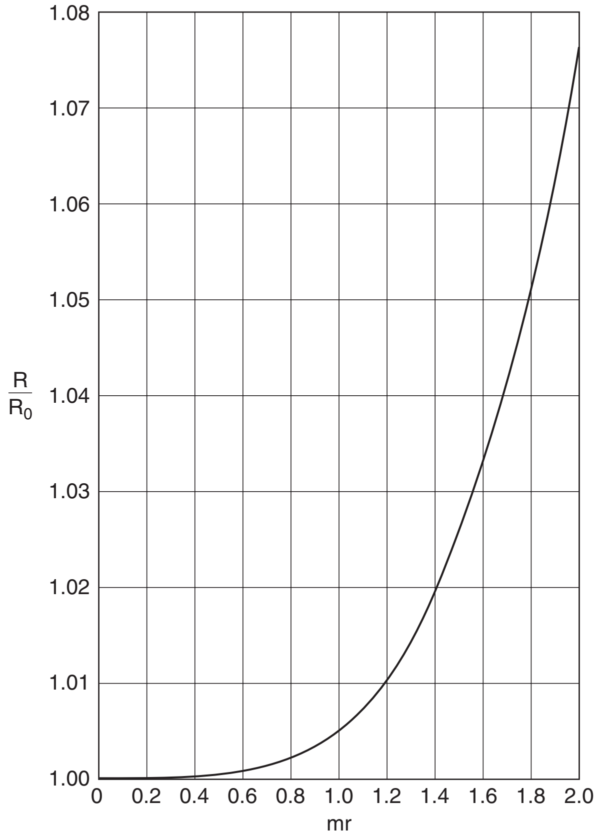


FIGURE 2-53 Ratio of A.C. to D.C. resistance of a cylindrical conductor. (*From W. D. Stevenson, Elements of Power Systems Analysis, New York, McGraw-Hill; 1962.*)

vol. 82, p. 537. These formulas have been compared with tests. The resistance ratio of square tubes is somewhat larger than that of round tubes. Values may be read from the curves of Fig. 4, Chap. 25, of *Electrical Coils and Conductors*.

Penetration Formula. For wires and tubes (and approximately for other compact shapes) where the resistance ratio is comparatively large, the conductor can be approximately considered to be replaced by its outer shell, of thickness equal to the "penetration depth," given by

$$\delta = \frac{1}{2\pi}\sqrt{\frac{10^7\rho}{f\mu_r}} \qquad \text{meters} \tag{2-156}$$

where ρ is the resistivity in ohm-meters, and μ_r is a presumed constant value of relative permeability. The resistance of the shell is then the effective resistance of the conductor. For Eq. (2-156) to be applicable, δ should be small compared with the dimensions of the cross section. In the case of tubes, δ is, evidently, less than the thickness of the tube. See Eq. (30), Chap. 19, of *Electrical Coils and Conductors*.

2.1.26 Electrostatics

Electrostatic Force. Electrically charged bodies exert forces on one another according to the following principles:

1. Like charged bodies repel; unlike charged bodies attract one another.
2. The force is proportional to the product of the magnitudes of the charges on the bodies.
3. The force is inversely proportional to the square of the distance between charges if the material in which the charges are immersed is extensive and possesses the same uniform properties in all directions.
4. The force acts along the line joining the centers of the charges.

Two concentrated charges Q_1 and Q_2 coulombs located R m apart experience a force between them of

$$F = \frac{Q_1 Q_2}{4\pi e_0 R^2} \qquad \text{newtons} \tag{2-157}$$

where $e_0 = 8.85419 \times 10^{-12}$ F/m and is the permittivity of free space.

Electrostatic Potential. The electric potential resulting from the location of charged bodies in the vicinity is called *electrostatic potential*. The potential at R m from a concentrated charge Q C is

$$\phi = \frac{Q}{4\pi e_0 R} \qquad \text{volts} \tag{2-158}$$

This potential is a scalar quantity.

Electric Field Intensity. The electric field intensity is the force per unit charge that would act at a point in the field on a very small test charge placed at that location. The electric field intensity E at a distance R m from a concentrated charge Q C is

$$E = \frac{Q}{4\pi e_0 R} \qquad \text{N/C} \tag{2-159}$$

Electric Potential Gradient in Electrostatic Fields. The space rate of change of the electric potential is the electric potential gradient of the field, symbolized by $\Delta\phi$. The general relationship between the gradient of the electric potential and the electric field intensity is

$$E = -\nabla\phi \qquad \text{V/m} \tag{2-160}$$

The units for the electric potential gradient, volts per meter, are frequently also used for the electric field intensity because their magnitudes are the same.

Electric Flux Density. The density of electric flux D in a region where simple dielectric materials exist is determined from the electric field intensity from

$$D = eE = e_0 KE \qquad \text{C/m}^2 \tag{2-161}$$

where K is a dimensionless number called the *dielectric constant*. In free space K is unity. For numerical values of dielectric constant of various dielectrics, see Sec. 4.

Polarization. The polarization is the excess of electric flux density that results in dielectric materials over that which would result at the same electric field intensity if the space were free of material substance. Thus

$$P = D - e_0 E \qquad \text{C/m}^2 \tag{2-162}$$

Crystalline Atomic Materials. In simple isotropic materials, the directions of the vectors P, D, and E are the same. For crystalline atomic structures that are not isotropic, Eq. (2-162) is the only relationship which is meaningful, and Eq. (2-161) should not be used.

Electric Flux. Electric flux and its density are related by

$$\psi = \int D \cos \alpha dA \qquad \text{coulombs} \tag{2-163}$$

where α is the angle between the direction of the electric flux density D and the normal at each differential surface area dA.

Capacitance. The capacitance between two oppositely charged bodies is the ratio of the magnitude of charge on either body to the difference of electric potential between them. Thus

$$C = \frac{Q}{V} \qquad \text{farads} \tag{2-164}$$

where Q is in coulombs and V is the voltage between the two equally but oppositely charged bodies, in volts.

Elastance. The reciprocal of capacitance, called *elastance*, is

$$S = V/Q \qquad \text{farads} \tag{2-165}$$

Electric Field Outside an Isolated Sphere in Free Space. The electric field intensity at a distance r m from the center of an isolated charged sphere located in free space is

$$E = \frac{Q}{4\pi e_0 r^2} \qquad \text{V/m} \tag{2-166}$$

where Q is the total charge (which is distributed uniformly) on the sphere.

Spherical Capacitor. The capacitance between two concentric charged spheres is

$$C = \frac{4\pi e_0 K}{1/R_1 - 1/R_2} \qquad \text{farads} \tag{2-167}$$

where R_1 is the outside radius of the inner sphere, R_2 is the inside radius of the outer sphere, and K is the dielectric constant of the space between them.

Electric Field Intensity Created by an Isolated, Charged, Long Cylindrical Wire in Free Space. The electric field intensity in the vicinity of a long, charged cylinder is

$$E = \frac{\Lambda}{2\pi e_0 r} \qquad \text{V/m} \tag{2-168}$$

where Λ is the charge per unit of length in coulombs per meter (distributed uniformly over the surface of the isolated cylinder) and r is the distance in meters from the center of the cylinder to the point at which the electric field intensity is evaluated.

Coaxial Cable. The capacitance per unit length of a coaxial cable composed of two concentric cylinders is

$$c = \frac{2\pi e_0 K}{\ln(R_2/R_1)} \qquad \text{F/m} \tag{2-169}$$

where R_1 is the outside radius of the inner cylinder, R_2 is the inside radius of the outer cylinder, and K is the dielectric constant of the space between the cylinders.

Two-Wire Line. The capacitance per unit length between two long, oppositely charged cylindrical conductors of equal radii, parallel and external to each other, is

$$c = \frac{2\pi e_0 K}{\ln\left[\frac{D}{2R} + \sqrt{\left(\frac{D}{2R}\right)^2 - 1}\right]} \qquad \text{F/m} \tag{2-170}$$

where D is the distance in meters between centers of the two cylindrical wires each with radius R and K is the uniform dielectric constant of all space external to the wires.

Capacitance of Two Flat, Parallel Conductors Separated by a Thin Dielectric. The capacitance is approximately

$$C = \frac{e_0 KA}{t} \qquad \text{farads} \tag{2-171}$$

where A is the area of either of the two conductors, t is the spacing between them, and K is the dielectric constant of the space between the conductors. Strictly, the linear dimensions of the flat conductors should be very large compared with the spacing between them. Good results are obtained from Eq. (2-171) even though the conductors are curved provided that the spacing t is small with respect to the radius of curvature.

Induced Charges. The surface of a conducting body, near a charge Q, through which no currents are flowing is an equipotential surface, a condition maintained by the motion of positive and negative charges to the parts of the conductor near Q and distant from it. Hence, the potential at any point on the conductor, due to all the charges of the system, is a constant. The charges on the conductors are said to be induced by Q, and the conductor is said to be electrified by induction.

Electrostatic Induction on Parallel Wires. Two insulated wires running parallel to a wire carrying a charge A C/m display a potential difference (provided that the two wires are not connected to each other or to other conductors) of

$$\phi = \frac{\Lambda}{2\pi e_0} \ln \frac{b}{a} \qquad \text{volts} \tag{2-172}$$

where b and a are the distances of the two insulated wires from the charged wire.

If the two wires are connected together, as, for example, through telephone instruments, the current flowing from one wire to the other is that required to equalize their potential difference.

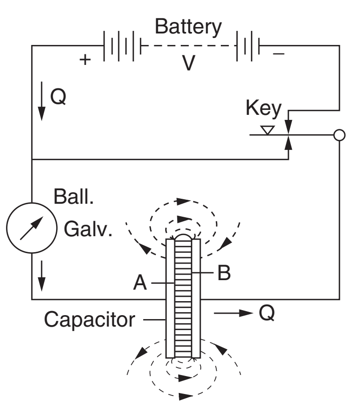


FIGURE 2-54 Circuit containing a capacitor.

2.1.27 The Dielectric Circuit

Circuit Concepts with Capacitive Elements. When a continuous voltage is applied to the terminals of a capacitor (AB, Fig. 2-54), a positive charge of electricity $+Q$ appears on one plate and a negative charge $-Q$ on the other. A quantity of electricity Q flows through the connecting wires, and this quantity of electricity is said to be displaced through the dielectric. An electrostatic field then exists between the two charged plates. Capacitors are introduced in Sec. 2.1.7. Capacitance formulas are given in those paragraphs.

Electrostatic Flux. The space between the plates of a capacitor can be treated as a dielectric circuit through which passes a dielectric flux ψ, in coulombs. In any dielectric circuit, one coulomb of electrostatic flux passes from each coulomb of positive charge to each coulomb of negative charge, and this is true with any insulating substance or group of substances. That is, electrostatic flux lines end only on charges of electricity. Their number is not affected when they pass from one dielectric to another, unless there is a charge of electricity on the surface of separation. Electrostatic flux lines are also called *lines of electrostatic induction*.

Capacitance to Neutral of a Conductor. The capacitance to neutral of a conductor in an AC line is defined as the capacitance that, when multiplied by $2\pi f$ and by the voltage to neutral, gives the charging current of the conductor, f being the frequency. This is not the same as the capacitance to a neutral wire measured electrostatically. The voltage to neutral of a single-phase line is one-half the voltage between conductors. The voltage to neutral of a balanced 3-phase line is equal to the voltage between conductors divided by 1.732.

When the conductors are round wires, for either single-phase or 3-phase overhead lines, the capacitance to neutral is

$$C = \frac{2\pi e_0 K}{\ln\sqrt{\frac{s}{d} + \left[\left(\frac{s}{d}\right)^2 - 1\right]^{1/2}}} \qquad \text{F/m, to neutral} \tag{2-173}$$

or, approximately,

$$C = \frac{0.0388}{\ln(2s/d)} \qquad \mu\text{F/mi, to neutral} \tag{2-174}$$

where s is the axial spacing and d is the diameter of the conductors, in the same units. Values of charging kVA for transmission lines are tabulated in Sec. 14.

The capacitance of a complete single-phase line is one-half the capacitance to neutral of one conductor. The capacitance of stranded conductors may be approximately calculated by using the outside diameter of the conductors. The capacitance of iron or steel conductors is calculated by the same formulas as that of copper conductors.

The preceding relations assume equilateral spacing for 3-phase systems. If unbalanced spacings are present and the phases are balanced by transposing the conductors over the length of the line, the approximate capacitance per phase can be obtained from the concepts of geometric mean distances D_m and D_s. Then, approximately,

$$C = \frac{2\pi e_0 k}{\ln(D_m D_s')} \qquad \text{f/m} \tag{2-175}$$

The self-geometric mean distance n in Eq. (2-175) differs slightly from that for D_s in Eq. (2-154), in that for good conductors, the transverse gradient of the electric field is confined principally to the airspace about the conductors, and D is slightly larger than D_s of Eq. (2-154). In the latter equation, internal flux linkages in the conductor contribute to the meaning of D_s.

Velocity of Propagation on Long Transmission Lines. The inductance and capacitive parameters per unit length of a transmission line determine the velocity with which such effects as switching surges are propagated along the line. The velocity of propagation is

$$\nu = \frac{1}{\sqrt{lc}} \qquad \text{m/s} \tag{2-176}$$

where l is the inductance per unit length from Eq. (2-154), and c is the capacitance per unit length from Eq. (2-175). Substituting these values gives

$$\nu = 3 \times 10^8 = \sqrt{\frac{\ln(D_m/D_s')}{K \ln(D_m/D_s)}} \qquad \text{m/s} \tag{2-177}$$

The fact that D is slightly larger than D_s produces a velocity of propagation along the transmission line which is slightly less than the velocity of propagation of electromagnetic radiation in free space (3 × 108 m/s). Since the dielectric constant K of the atmosphere surrounding the transmission line may be somewhat greater than unity, the velocity of propagation may be reduced slightly more. Magnetic materials in the conductors tend to increase the inductance in the denominator of Eq. (2-176) and reduce further the velocity of propagation by a small amount.

Dielectric Strength of Insulating Materials. The dielectric strength of insulating materials (rupturing voltage gradient) is the maximum voltage per unit thickness that a dielectric can withstand in a uniform field before it breaks down electrically. The dielectric strength is usually measured in kilovolts per millimeter or per inch. It is necessary to define the dielectric strength in terms of a uniform field, for instance, between large parallel plates a short distance apart. If the striking voltage is determined between two spheres or electrodes of other defined shape, this fact must be stated. In designing insulation, a factor of safety is assumed depending upon conditions of operation. For numerical values of rupturing voltage gradients of various insulating materials, see Sec. 4.

2.1.28 Dielectric Loss and Corona

Dielectric Hysteresis and Conductance. When an alternating voltage is applied to the terminals of a capacitor, the dielectric is subjected to periodic stresses and displacements. If the material were perfectly elastic, no energy would be lost during any cycle, because the energy stored during the periods of increased voltage would be given up to the circuit when the voltage is decreased. However, since the electric elasticity of dielectrics is not perfect, the applied voltage has to overcome molecular friction or viscosity, in addition to the elastic forces. The work done against friction is converted into heat and is lost. This phenomenon resembles magnetic hysteresis (Sec. 2.1.23) in some respects but differs in others. It has commonly been called *dielectric hysteresis* but is now often called *dielectric loss*. The energy lost per cycle is proportional to the square of the applied voltage. Methods of measuring dielectric loss are described in Sec. 3.

An imperfect capacitor does not return on discharge the full amount of energy put into it. Sometime after the discharge, an additional discharge may be obtained. This phenomenon is known as *dielectric absorption*.

A capacitor that shows such a loss of power can be replaced for purposes of calculation by a perfect capacitor with an ohmic conductance shunted around it. This conductance (or "leakance") is of such value that its *PR* loss is equal to the loss of power from all causes in the imperfect capacitor. The actual current through the capacitor is then considered as consisting of two components—the leading reactive component through the ideal capacitor and the loss component, in phase with the voltage, through the shunted conductance.

Electrostatic Corona. When the electrostatic flux density in the air exceeds a certain value, a discharge of pale violet color appears near the adjacent metal surfaces. This discharge is called *electrostatic corona*. In the regions where the corona appears, the air is electrically ionized and is a conductor of electricity. When the voltage is raised further, a brush discharge takes place, until the whole thickness of the dielectric is broken down and a disruptive discharge, or spark, jumps from one electrode to the other.

Corona involves power loss, which may be serious in some cases, as on transmission lines (Secs. 14 and 15). Corona can form at sharp corners of high-voltage switches, bus bars, etc., so the radii of such parts are made large enough to prevent this. A voltage of 12 to 25 kV between conductors separated by a fraction of an inch, as between the winding and core of a generator or between sections of the winding of an air-blast transformer, can produce a voltage gradient sufficient to cause corona.

A voltage of 100 to 200 kV may be required to produce corona on transmission-line conductors that are separated by several feet. Corona can have an injurious effect on fibrous insulation. For numerical data in application to transmission lines see Secs. 14 and 15.

BIBLIOGRAPHY

Abraham, M.: The *Classical Theory of Electricity and Magnetism*, revised by R. Becker, translated into English by J. Dougall. Glasgow, Blackie & Son, Ltd., 1932.

Anderson, P. M.: *Analysis of Faulted Power Systems.* Ames, I.A., Iowa State University Press, 1973.

Balabanian, N., Bickart, T. A., and Seshu, S.: *Electrical Network Theory*. New York, John Wiley & Sons, Inc., 1969.

Bergen, A. R.: *Power Systems Analysis*. Englewood Cliffs, N.J., Prentice-Hall, 1986.

Bitter F.: *Introduction to Ferromagnetism*. New York, McGraw-Hill, 1937.

Boast, W. B.: *Vector Fields*. New York, Harper & Row Publishers, Inc., 1964.

Brittain, J. E. (Ed.): *Turning Points in American Electrical History*. New York, IEEE Press, 1977.

Bush, V.: *Operational Circuit Analysis*. New York, John Wiley & Sons, Inc., 1937.

Clarke, E.: *Circuit Analysis of A-C Power Systems*. New York, John Wiley & Sons, Inc., 1943, vol. I, and 1950, vol. II.

Dwight, H. B.: *Electrical Coils and Conductors*. New York, McGraw-Hill, 1945.

Dwight, H. B.: *Electrical Elements of Power Transmission Lines*. New York, The Macmillan Company, 1954.

Electrical Transmission and Distribution Reference Book, 4th ed. Westinghouse Electric Corporation, 1950.

Encyclopedia of Physics, 3d ed. New York, Van Nostrand Reinhold Co., 1985.

Faraday, M.: *Experimental Researches in Electricity*, 3 vols. London, B. Quaritch, 1839–1855.

Fitzgerald, A. E., and Kingsley, C., Jr.: *Electric Machinery*, 3d ed. New York, McGraw-Hill, 1961.

Frank, N. H.: *Introduction to Electricity and Optics*, 2d ed. New York, McGraw-Hill, 1950.

Gardner, M. F., and Barnes, J. L.: *Transients in Linear Systems*. New York, John Wiley & Sons, Inc., 1942.

Ham, I. M., and Slemon, G. R.: *Scientific Basis of Electrical Engineering*. New York, John Wiley and Sons, Inc., 1961.

Harnwell, G. P.: *Principles of Electricity and Electromagnetism*, 2d ed. New York, McGraw-Hill, 1949.

Hayt, W. H.: *Engineering Electromagnetics*, 4th ed. New York, McGraw-Hill, 1984.

Hayt, W. H., and Kemmerly, I. E.: *Engineering Circuit Analysis*, 5th ed. New York, McGraw-Hill, 1993.

Huelsman, L.: *Circuits, Matrices, and Linear Vector Spaces*. New York, McGraw-Hill, 1963.

Jeans, I. H.: *Mathematical Theory of Electricity and Magnetism*. New York, Cambridge University Press, 1908.

Keown, I. L.: *PSPICE and Circuit Analysis*. New York, Merrill Publishing, 1991.

Kusic, G. L.: *Computer-Aided Power Systems Analysis*. Englewood Cliffs, N.J., Prentice-Hall, 1986.

Lee, R., Wilson, L., and Carter, C. E.: *Electronic Transformers and Circuits*, 3d ed. New York, John Wiley & Sons, Inc., 1988.

Maxwell, I. C.: *A Treatise on Electricity and Magnetism,* 2 vols. New York, Oxford University Press, 1904.

MIT Staff: *Magnetic Circuits and Transformers.* New York, John Wiley & Sons, Inc., 1943.

Nilsson, J. W.: *Electric Circuits,* 5th ed. Reading, Mass., Addison-Wesley Publishing Company, 1996.

Page, L., and Adams, N. I.: *Principles of Electricity*. Princeton, N.J., D. Van Nostrand Company, Inc., 1934.

Peek, F. W., Jr.: *Dielectric Phenomena in High-Voltage Engineering.* New York, McGraw-Hill, 1929.

Rashid, M. H.: *SPICE for Circuits and Electronics Using PSPICE.* Englewood Cliffs, N.J., Prentice-Hall, 1990.

Rosa, E. B., and Grover, F. W.: Formulas and Tables for the Calculation of Mutual and Self-Inductance, *NBS Sci. Paper* 169, 1916. Published also as Pt. I of vol. 8, *NBS Bull.* Contains also skin-effect tables.

Ryder, J. D., and Fink, D. G.: *Engineers and Electrons*. New York, IEEE Press, 1984.

Sears, F. W.: *Principles of Physics*. Reading, Mass., Addison-Wesley Publishing Company, Inc., 1946, vol. 2, *Electricity and Magnetism*.

Seshu, S., and Balabanian, N.: *Linear Network Analysis*. New York, John Wiley & Sons, Inc., 1959.

Skilling, H. H.: *Electromechanics*. New York, John Wiley & Sons, Inc., 1962.

Smythe, W. R.: *Static and Dynamic Electricity*, 3d ed. New York, McGraw-Hill, 1967.

Stevenson, W. D., Jr.: *Elements of Power System Analysis*, 4th ed. New York, McGraw-Hill, 1982.

Thorpe, T. W.: *Computerized Circuit Analysis with SPICE: A Complete Guide to SPICE, with Applications.* New York, John Wiley & Sons, Inc., 1992.

Tuinenga, P. W.: *SPICE, A Guide to Circuit Simulation and Analysis Using PSPICE*. Englewood Cliffs, N.J., Prentice-Hall, 1988.

Van Valkenburg, M. E.: *Circuit Theory: Foundations and Classical Contributions.* Stroudsburg, PA., Dowden, Hutchinson, and Ross, 1974.

Van Valkenburg, M. E.: *Linear Circuits.* Englewood Cliffs, N.J., Prentice-Hall, 1982.

Van Valkenburg, M. E.: *Network Analysis*. Englewood Cliffs, N.J., Prentice-Hall, 1974.

Wildi, T.: *Electric Power Technology*. New York, John Wiley & Sons, Inc., 1981.

Woodruff, L. F.: *Principles of Electric Power Transmission*. New York, John Wiley & Sons, Inc., 1938.

Internet References

http://www.lectureonline.cl.msu.edu/~mmp/applist/induct/faraday.htm

http://www.ee.byu.edu/em/amplaw2.htm.

http://www.ee.byu.edu/ee/em/eleclaw.htm

Software References

MathCAD 11.2A, © 1986–2003 Mathsoft Engineering & Education, Inc.

OrCAD Capture 9.1, © 1985–1999 OrCAD, Inc.

SECTION 3
MEASUREMENTS AND INSTRUMENTS*

Gerald J. Fitzpatsick
Project Leader, Advanced Power System Measurements
National Institute of Standard and Technology

CONTENTS

3.1 ELECTRIC AND MAGNETIC MEASUREMENTS

3.1.1 General

Measurement of a quantity consists either of its comparison with a unit quantity of the same kind or of its determination as a function of quantities of different kinds whose units are related to it by known physical laws. An example of the first kind of measurement is the evaluation of a resistance

*Grateful acknowledgement is given to Norman Belecki, George Burns, Forest Harris, and B.W. Mangum for most of the material in this section.

(in ohms) with a Wheatstone bridge in terms of a calibrated resistance and a ratio. An example of the second kind is the calibration of the scale of a wattmeter (in watts) as the product of current (in amperes) in its field coils and the potential difference (in volts) impressed on its potential circuit.

The units used in electrical measurements are related to the metric system of mechanical units in such a way that the electrical units of power and energy are identical with the corresponding mechanical units. In 1960, the name *Système International* (abbreviated SI), now in use throughout the world, was assigned to the system based on the meter-kilogram-second-ampere (abbreviated mksa). The mksa units are identical in value with the practical units—volt, ampere, ohm, coulomb, farad, henry—used by engineers. Certain prefixes have been adopted internationally to indicate decimal multiples and fractions of the basic units.

A *reference standard* is a concrete representation of a unit or of some fraction or multiple of it having an assigned value which serves as a measurement base. Its assignment should be traceable through a chain of measurements to the National Reference Standard maintained by the National Institute of Standards and Technology (NIST). Standard cells and certain fixed resistors, capacitors, and inductors of high quality are used as reference standards.

The National Reference Standards maintained by the NIST comprise the legal base for measurements in the United States. Other nations have similar laboratories to maintain the standards which serve as their measurement base. An international bureau—Bureau International des Poids et Mesures (abbreviated BIPM) in Sèvres, France—also maintains reference standards and compares standards from the various national laboratories to detect and reconcile any differences that might develop between the as-maintained units of different countries.

At NIST, the reference standard of resistance is a group of 1-Ω resistors, fully annealed and mounted strain-free out of contact with the air, in sealed containers. The reference standard of capacitance is a group of 10-pF fused-silica-dielectric capacitors whose values are assigned in terms of the calculable capacitor used in the ohm determination. The reference standard of voltage is a group of standard cells continuously maintained at a constant temperature.

The "absolute" experiments from which the value of an electrical unit is derived are measurements in which the electrical unit is related directly to appropriate mechanical units. In recent *ohm* determinations, the value of a capacitor of special design was calculated from its measured dimensions, and its impedance at a known frequency was compared with the resistance of a special resistor. Thus, the ohm was assigned in terms of length and time. The as-maintained ohm is believed to be within 1 ppm of the defined SI unit. Recent *ampere* determinations, used to assign the *volt* in terms of current and resistance, derived the ampere by measuring the force between current-carrying coils of a mutual inductor of special construction whose value was calculated from its measured dimensions. The voltage drop of this current in a known resistor was used to assign the emf of the standard cells which maintain the volt. The stated uncertainty of these ampere determinations ranges from 4 to 7 ppm, and the departure of value of the "legal" volt from the defined SI unit carries the same uncertainty. Since 1972, the assigned emf of the standard cells in the reference group which maintains the legal volt is monitored (and reassigned as necessary) in terms of atomic constants (the ratio of Planck's constant to electron charge) and a microwave frequency by an ac Josephson experiment in which their voltage is measured with respect to the voltage developed across the barrier junction between two superconductors irradiated by microwave energy and biased with a direct current. This experiment appears to be repeatable within 0.1 ppm. It should be noted that while the Josephson experiment may be used to maintain the legal volt at a constant level, it is not used to define the SI unit.

Precision—a measure of the spread of repeated determinations of a particular quantity—depends on various factors. Among these are the resolution of the method used, variations in ambient conditions (such as temperature and humidity) that may influence the value of the quantity or of the reference standard, instability of some element of the measuring system, and many others. In the National Laboratory of the National Institute of Standards and Technology, where every precaution is taken to obtain the best possible value, intercomparisons may have a precision of a few parts in 10^7. In commercial laboratories, where the objective is to obtain results that are reliable but only to the extent justified by engineering or other requirements, precision ranges from this figure to a part in 10^3 or more, depending on circumstances. For commercial measurements such as the sale of electrical energy, where the cost of measurement is a critical factor, a precision of 1 or 2% is considered acceptable in some jurisdictions.

The use of digital instruments occasionally creates a problem in the evaluation of precision, that is, all results of a repeated measurement may be identical due to the combination of limited resolution and quantized nature of the data. In these cases, the least count and sensitivity of the instrumentation must be taken into account in determining precision.

Accuracy—a statement of the limits which bound the departure of a measured value from the true value of a quantity—includes the imprecision of the measurement, together with all the accumulated errors in the measurement chain extending downward from the basic reference standards to the specific measurement in question. In engineering measurement practice, accuracies are generally stated in terms of the values assigned to the National Reference Standards—the *legal* units. It is only rarely that one needs also to state accuracy in terms of the defined SI unit by taking into account the uncertainty in the assignment of the National Reference Standard.

General precautions should be observed in electrical measurements, and sources of error should be avoided, as detailed below:

1. The accuracy limits of the instruments, standards, and methods used should be known so that appropriate choice of these measuring elements may be made. It should be noted that instrument *accuracy classes* state the "initial" accuracy. Operation of an instrument, with energy applied over a prolonged period, may cause errors due to elastic fatigue of control springs or resistance changes in instrument elements because of heating under load. ANSI C39.1 specifies permissible limits of error of portable instruments because of sustained operation.
2. In any other than rough determinations, the *average of several readings* is better than one. Moreover, the alteration of measurement conditions or techniques, where feasible, may help to avoid or minimize the effects of accidental and systematic errors.
3. The *range* of the measuring instrument should be such that the measured quantity produces a reading large enough to yield the desired precision. The deflection of a measuring instrument should preferably exceed half scale. Voltage transformers, wattmeters, and watthour meters should be operated near to rated voltage for best performance. Care should be taken to avoid either momentary or sustained overloads.
4. *Magnetic fields*, produced by currents in conductors or by various classes of electrical machinery or apparatus, may combine with the fields of portable instruments to produce errors. Alternating or time-varying fields may induce emfs in loops formed in connections or the internal wiring of bridges, potentiometers, etc. to produce an error signal or even "electrical noise" that may obscure the desired reading. The effects of stray alternating fields on ac indicating instruments may be eliminated generally by using the average of readings taken with direct and reversed connections; with direct fields and dc instruments, the second reading (to be averaged with the first) may be taken after rotating the instrument through 180°. If instruments are to be mounted in magnetic panels, they should be calibrated in a panel of the same material and thickness. It also should be noted that Zener-diode-based references are affected by magnetic fields. This may alter the performance of digital meters.
5. In measurements involving high resistances and small currents, *leakage paths* across insulating components of the measuring arrangement should be eliminated if they shunt portions of the measuring circuit. This is done by providing a guard circuit to intercept current in such shunt paths or to keep points at the same potential between which there might otherwise be improper currents.
6. Variations in *ambient temperature* or internal temperature rise from self-heating under load may cause errors in instrument indications. If the temperature coefficient and the instrument temperature are known, readings can be corrected where precision requirements justify it. Where measurements involve extremely small potential differences, thermal emfs resulting from temperature differences between junctions of dissimilar metals may produce errors; heat from the observer's hand or heat generated by the friction of a sliding contact may cause such effects.
7. *Phase-defect angles* in resistors, inductors, or capacitors and in instruments and instrument transformers must be taken into account in many ac measurements.
8. Large *potential differences* are to be avoided between the windings of an instrument or between its windings and frame. Electrostatic forces may produce reading errors, and very large potential

difference may result in insulating breakdown. Instruments should be connected in the ground leg of a circuit where feasible. The moving-coil end of the voltage circuit of a wattmeter should be connected to the same line as the current coil. When an instrument must be at a high potential, its case must be adequately insulated from ground and connected to the line in which the instrument circuit is connected, or the instrument should be enclosed in a screen that is connected to the line. Such an arrangement may involve shock hazard to the operator, and proper safety precautions must be taken.

9. *Electrostatic charges* and consequent disturbance to readings may result from rubbing the insulating case or window of an instrument with a dry dustcloth; such charges can generally be dissipated by breathing on the case or window. Low-level measurements in very dry weather may be seriously affected by charges on the clothing of the observer; some of the synthetic textile fibers—such as nylon and Dacron—are particularly strong sources of charge; the only effective remedy is the complete screening of the instrument on which charges are induced.

10. *Position influence* (resulting from mechanical unbalance) may affect the reading of an analog-type indicating instrument if it is used in a position other than that in which it was calibrated. Portable instruments of the better accuracy classes (with antiparallax mirrors) are normally intended to be used with the axis of the moving system vertical, and the calibration is generally made with the instrument in this position.

3.1.2 Detectors and Galvanometers

Detectors are used to indicate approach to balance in bridge or potentiometer networks. They are generally responsive to small currents or voltages, and their sensitivity—the value of current or voltage that will produce an observable indication—ultimately limits the resolution of the network as a means for measuring some electrical quantity.

Galvanometers are deflecting instruments which are used, mainly, to *detect* the presence of a small electrical quantity—current, voltage, or charge—but which are also used in some instances to measure the quantity through the magnitude of the deflection.

The D'Arsonval (moving-coil) galvanometer consists of a coil of fine wire suspended between the poles of a permanent magnet. The coil is usually suspended from a flat metal strip which both conducts current to it and provides control torque directed toward its neutral (zero-current) position. Current may be conducted from the coil by a helix of fine wire which contributes very little to the control torque (pendulous suspension) or by a second flat metal strip which contributes significantly to the control torque (taut-band suspension). An iron core is usually mounted in the central space enclosed by the coil, and the pole pieces of the magnet are shaped to produce a uniform radial field throughout the space in which the coil moves. A mirror attached to the coil is used in conjunction with a lamp and scale or a telescope and scale to indicate coil position.

The pendulous-suspension type of galvanometer has the advantage of higher sensitivity (weaker control torque) for a suspension of given dimensions and material and the disadvantage of responsiveness to mechanical disturbances to its supporting platform, which produce anomalous motions of the coil. The taut-suspension type is generally less sensitive (stiffer control torque) but may be made much less responsive to mechanical disturbances if it is properly balanced, that is, if the center of mass of the moving system is in the axis of rotation determined by the taut upper and lower suspensions.

Galvanometer sensitivity can be expressed in a number of ways, depending on application:

1. The *current* constant is the current in microamperes that will produce unit deflection on the scale—usually a deflection of 1 mm on a scale 1 m distant from the galvanometer mirror.
2. The *megohm* constant is the number of megohms in series with the galvanometer through which 1 V will produce unit deflection. It is the reciprocal of the current constant.
3. The *voltage* constant is the number of microvolts which, in a critically damped circuit (or another specified damping), will produce unit deflection.

4. The *coulomb* constant is the charge in microcoulombs which, at a specified damping, will produce unit ballistic throw.

5. The *flux-linkage* constant is the product of change of induction and turns of the linking search coil which will produce unit ballistic throw.

All these sensitivities (galvanometer response characteristics) can be expressed in terms of current sensitivity, circuit resistance in which the galvanometer operates, relative damping, and period. If we define *current* sensitivity S_i as deflection per unit current, then—in appropriate units—the *voltage* sensitivity (the deflection per unit voltage) is

$$S_e = \frac{S_i}{R}$$

where R is the resistance of the circuit, including the resistance of the galvanometer coil. The *coulomb* sensitivity is

$$\frac{\theta}{Q} = \frac{2\pi}{T_o} S_i \exp\left(\frac{-\gamma}{\sqrt{1-\gamma^2}} \tan^{-1} \frac{\sqrt{1-\gamma^2}}{\gamma}\right)$$

where T_o is the undamped period and γ is the relative damping in the operating circuit. The *flux-linkage* sensitivity is

$$\frac{\theta}{\int e\,dt} \approx S_i \frac{2\pi}{T_o} \frac{1}{2R_c} \frac{1}{1-\gamma_0}$$

for the case of greatest interest—maximum ballistic response—where the galvanometer is heavily overdamped, γ_0 being the open-circuit relative damping, $\int e\,dt$ the time integral of induced voltage or the change in flux linkages in the circuit, and R_c the circuit resistance (including that of the galvanometer) for which the galvanometer is critically damped.

Galvanometer motion is described by the differential equation

$$P\ddot{\theta} + \left(K + \frac{G^2}{R}\right)\dot{\theta} + U\theta = \frac{GE}{R}$$

where θ is the angle of deflection in radians, P is the moment of inertia, K is the mechanical damping coefficient, G is the motor constant (G = coil area turns × air-gap field), R is total circuit resistance (including the galvanometer), and U is the suspension stiffness. If the viscous and circuital damping are combined,

$$K + G^2/R = A$$

the roots of the auxiliary equation are

$$m = \frac{A}{2P} \pm \sqrt{\frac{A^2}{4P^2} - \frac{U}{P}}$$

Three types of motion can be distinguished.

1. *Critically damped* motion occurs when $A^2/4P^2 = U/P$. It is an aperiodic, or deadbeat, motion in which the moving system approaches its equilibrium position without passing through it in the shortest time of any possible aperiodic motion. This motion is described by the equation

$$y = 1 - \left(1 + \frac{2\pi t}{T_o}\right)\exp\left(\frac{-2\pi t}{T_o}\right)$$

where y is the fraction of equilibrium deflection at time t and T_o is the undamped period of the galvanometer—the period that the galvanometer would have if $A = 0$. If the total damping coefficient at critical damping is A_c, we can define relative damping as the ratio of the damping coefficient A for a specific circuit resistance to the value A_c it has for critical damping—$\gamma = A/A_c$, which is unity for critically damped motion.

2. In *overdamped* motion, the moving system approaches its equilibrium position without overshoot and more slowly than in critically damped motion. This occurs when

$$\frac{A^2}{4P^2} > \frac{U}{P}$$

and $\gamma > 1$. For this case, the motion is described by the equation

$$y = 1 - \left(\frac{\gamma}{\sqrt{\gamma^2 - 1}} \sinh \frac{2\pi t}{T_o} \sqrt{\gamma^2 - 1} + \cosh \frac{2\pi t}{T_o} \sqrt{\gamma^2 - 1}\right) \exp\left(\frac{-2\pi t}{T_o} \gamma\right)$$

3. In *underdamped* motion, the equilibrium position is approached through a series of diminishing oscillations, their decay being exponential. This occurs when

$$\frac{A^2}{4P^2} > \frac{U}{P}$$

and $\gamma < 1$. For this case, the motion is described by the equation

$$y = 1 - \frac{1}{\sqrt{1 - \gamma^2}} \left[\sin\left(\frac{2\pi t}{T_o} \sqrt{1 - \gamma^2} + \sin^{-1} \sqrt{1 - \gamma^2}\right)\right] \exp\left(\frac{-2\pi t}{T_o} \gamma\right)$$

Damping factor is the ratio of deviations of the moving system from its equilibrium position in successive swings. More conveniently, it is the ratio of the equilibrium deflection to the "overshoot" of the first swing past the equilibrium position, or

$$F = \frac{\theta_1 - \theta_F}{\theta_F - \theta_2} = \frac{\theta_F}{\theta_1 - \theta_F}$$

where θ_F is the equilibrium deflection and θ_1 and θ_2 are the first maximum and minimum deflections of the damped system. It can be shown that damping factor is connected to relative damping by the equation

$$F = \exp\left(\frac{\pi\gamma}{\sqrt{1 - \gamma^2}}\right)$$

The *logarithmic decrement* of a damped harmonic motion is the naperian logarithm of the ratio of successive swings of the oscillating system. It is expressed by the equation

$$\ln \frac{\theta_1 - \theta_F}{\theta_F - \theta_2} = \ln \frac{\theta_F}{\theta_1 - \theta_F} = \lambda$$

and in terms of relative damping

$$\lambda = \frac{\pi\gamma}{\sqrt{1 - \gamma^2}}$$

The *period* of a galvanometer (and, generally, of any damped harmonic oscillator) can be stated in terms of its undamped period T_o and its relative damping γ as $T = T_o/\sqrt{1 - \gamma^2}$.

Reading time is the time required, after a change in the quantity measured, for the indication to come and remain within a specified percentage of its final value. Minimum reading time depends on the relative damping and on the required accuracy (Table 3-1). Thus, for a reading within 1% of equilibrium value, minimum time will be required at a relative damping of $\gamma = 0.83$. Generally in indicating instruments, this is known as *response time* when the specified accuracy is the stated accuracy limit of the instrument.

TABLE 3-1 Minimum Reading Time for Various Accuracies

Accuracy, percent	Relative damping	Reading time/free period
10	0.6	0.37
1	0.83	0.67
0.1	0.91	1.0

External critical damping resistance (CDRX) is the external resistance connected across the galvanometer terminals that produces critical damping ($\gamma = 1$).

Measurement of damping and its relation to circuit resistance can be accomplished by a simple procedure in the circuit of Fig. 3-1. Let R_a be very large (say, 150 kΩ) and R_b small (say, 1 Ω) so that when E is a 1.5-V dry cell, the driving voltage in the local galvanometer loop is a few microvolts (say, 10 μV). Since circuital damping is related to *total* circuit resistance ($R_c + R_b + R_g$), the galvanometer resistance R_g must be determined first. If R_c is adjusted to a value that gives a convenient deflection and then to a new value R_c' for which the deflection is cut in half, we have $R_g = R_c' - 2R_c - R_b$. Now, let R_c be set at such a value that when the switch is closed, the overshoot is readily observed. After noting the open-circuit deflection θ_o, the switch is closed and the peak value θ, of the first overswing, and the final deflection θ_F are noted. Then

$$\ln \frac{\theta_F - \theta_o}{\theta_1 - \theta_F} = \frac{\pi \gamma_1}{\sqrt{1 - \gamma_1^2}}$$

γ_1 being the relative damping corresponding to the circuit resistance $R_1 = R_g + R_b + R_c$. The switch is now opened, and the first overswing θ_2 past the open-circuit equilibrium position θ_o is noted. Then

$$\ln \frac{\theta_F - \theta_o}{\theta_2 - \theta_o} = \frac{\pi \gamma_o}{\sqrt{1 - \gamma_o^2}}$$

γ_o being the open-circuit relative damping. The relative damping γ_x for any circuit resistance R_x is given by the relation

$$\frac{R_x}{R_1} = \frac{\gamma_1 - \gamma_o}{\gamma_x - \gamma_o}$$

where it should be noted that the galvanometer resistance R_g is included in both R_x and R_1. For critical damping R_d can be computed by setting $\gamma_x = 1$, and the external critical damping resistance $CDRX = R_d - R_g$.

Galvanometer shunts are used to reduce the response of the galvanometer to a signal. However, in any sensitivity-reduction network, it is important that relative damping be preserved for proper operation. This can always be achieved by a suitable combination of series and parallel resistance. In Fig. 3-2, let r be the external circuit resistance and R_g the galvanometer resistance such that $r + R_g$ gives an acceptable damping (e.g., $\gamma = 0.8$) at maximum sensitivity. This damping will be preserved when the sensitivity-reduction network (S, P) is inserted, if $S = (n - 1)r$ and $P = nr/(n - 1)$, n being the factor by which response is to be reduced. The Ayrton-Mather shunt, shown

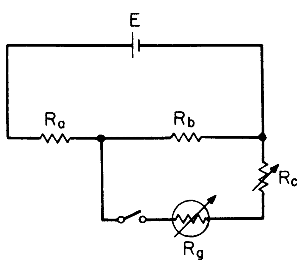


FIGURE 3-1 Determination of relative damping.

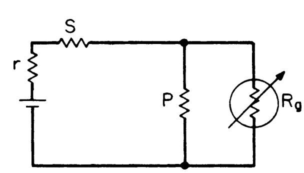


FIGURE 3-2 Galvanometer shunt.

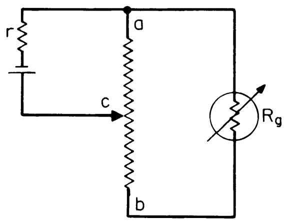


FIGURE 3-3 Ayrton-Mather universal shunt.

in Fig. 3-3, may be used where the circuit resistance r is so high that it exerts no appreciable damping on the galvanometer. R_{ab} should be such that correct damping is achieved by $R_{ab} + R_g$. In this network, sensitivity reduction is

$$n = R_{ac}/R_{ab}$$

and the ratio of galvanometer current I_g to line current I is

$$\frac{I_g}{I} = \frac{R_{ab}}{n(R_g + R_{ab})}$$

The *ultimate resolution* of a detection system is the magnitude of the signal it can discriminate against the noise background present. In the absence of other noise sources, this limit is set by the *Johnson noise* generated by electron thermal agitation in the resistance of the circuit. This is expressed by the formula $e = \sqrt{4k\theta Rf}$, where e is the rms noise voltage developed across the resistance R, k is Boltzmann's constant 1.4×10^{-23} J/K, θ is the absolute temperature of the resistor in kelvin, and f is the bandwidth over which the noise voltage is observed. At room temperature (300 K) and with the assumption that the peak-to-peak voltage is 5 × rms value, the peak-to-peak Johnson noise voltage is $6.5 \times 10^{-10} \sqrt{Rf}$ V. If, in a dc system, we use the approximation that $f = 1/3t$, where t is the system's response time, the Johnson voltage is $4 \times 10^{-10} \sqrt{R/t}$V (peak to peak).

By using reasonable approximations, it can be shown that the random brownian-motion deflections of the moving system of a galvanometer, arising from impulses by the molecules in the air around it, are equivalent to a voltage indication $e = 5 \times 10^{-10} \sqrt{R/T}V$ (peak to peak), where R is circuit resistance and T is the galvanometer period in seconds. If the galvanometer damping is such that its response time is $t = 2T/3$ (for $\gamma \approx 0.8$), the Johnson noise voltage to which it responds is about $5 \times 10^{-10} \sqrt{R/t}$V (peak to peak). This value represents the limiting resolution of a galvanometer, since its response to smaller signals would be obscured by the random excursions of its moving system. Thus, a galvanometer with a 4-s-period would have a limiting resolution of about 2 nV in a 100-Ω circuit and 1 nV in a 25-Ω circuit.

It is not surprising that one arrives at the same value from considerations either of random electron motions in the conductors of the measuring circuit or of molecular motions in the fluid that surrounds the system. The resulting figure rests on the premise that the law of equipartition of energy applies to the measuring system and that the galvanometer coil—a body with one degree of freedom—is statically in thermal equilibrium with its surroundings.

Optical systems used with galvanometers and other indicating instruments avoid the necessity for a mechanical pointer and thus permit smaller, simpler balancing arrangements because the mirror attached to the moving system can be symmetrically disposed close to the axis of rotation. In portable instruments, the entire system—source, lenses, mirror, scale—is generally integral with the instrument, and the optical "pointer" may be folded one or more times by fixed mirrors so that it is actually much longer than the mechanical dimensions of the instrument case. In some instances, the angular displacement may be magnified by use of a cylindrical lens or mirror. For a wall- or bracket-mounted galvanometer, the lamp and scale arrangement is external, and the length of the light-beam pointer can be controlled. Whatever the arrangement, the pointer length cannot be indefinitely extended with consequent increase in resolution at the scale. The optical resolution of such a system is, in any event, limited by image diffraction, and this limit—for a system limited by a circular aperture—is $\alpha \approx 1.2\lambda/nd$, where α is the angle subtended by resolvable points, λ is the wavelength of the light, n is the index of refraction of the image space, and d is the aperture diameter. In this case, d is the diameter of the moving-system mirror, and $n = 1$ for air. If we assume that points 0.1 mm apart can just be resolved by the eye at normal reading distance, the resolution limit is reached at a scale distance of about 2 m in a system with a 1-cm mirror, which uses no optical magnification. Thus, for the usual galvanometer, there is no profit in using a mirror-scale separation greater than 2 m. Since resolution is a matter of subtended angle, the corresponding scale distance is proportionately less for systems that make use of magnification.

The *photoelectric galvanometer amplifier* is a detector system in which the light beam from the moving-system mirror is split between two photovoltaic cells connected in opposition, as shown

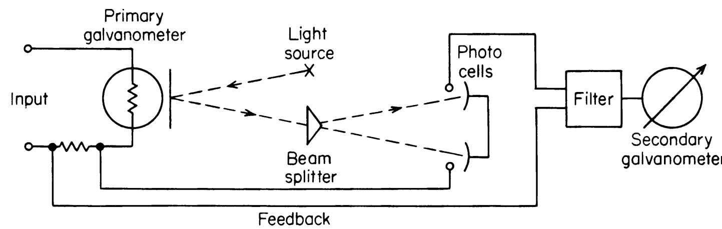


FIGURE 3-4 Photoelectric galvanometer amplifier.

in Fig. 3-4. As the mirror of the primary galvanometer turns in response to an input signal, the light flux is increased on one of the photocells and decreased on the other, resulting in a current and thence an enhanced signal in the circuit of the secondary (reading) galvanometer. Since the photocells respond to the total light flux on their sensitive elements, the system is not subject to resolution limitation by diffraction as is the human eye, and the ultimate resolution of the primary instrument—limited only by its brownian motion and the Johnson noise of the input circuit—may be realized.

Electronic instruments for low-level dc signal detection are more convenient, more rugged, and less susceptible to mechanical disturbances than is a galvanometer. However, considerable filtering, shielding, and guarding must be used to minimize electrical interference and noise. On the other hand, a galvanometer is an extremely efficient low-pass filter, and when operated to make optimal use of its design characteristics, it is still the most sensitive low-level dc detector. Electronic detectors generally make use of either a mechanical or a transistor chopper driven by an oscillator whose frequency is chosen to avoid the local power frequency and its harmonics. This modulator converts the dc input signal to ac, which is then amplified, demodulated, and displayed on an analog-type indicating instrument or fed to a recording device or a signal processor.

AC detectors used for balancing bridge networks are usually tuned low-level amplifiers coupled to an appropriate display device. The narrower the passband of the amplifier, the better the signal resolution, since the narrow passband discriminates against noise of random frequency in the input circuit. Adjustable-frequency amplifier-detectors basically incorporate a low-noise preamplifier followed by a high-gain amplifier around which is a tunable feedback loop whose circuit has zero transmission at the selected frequency so that the negative-feedback circuit controls the overall transfer function and acts to suppress signals except at the selected frequency. The amplifier output may be rectified and displayed on a dc indicating instrument, and added resolution is gained by introducing phase selection at the demodulator, since the wanted signal is regular in phase, while interfering noise is generally random. In detectors of this type, in phase and quadrature signals can be displayed separately, permitting independent balancing of bridge components. Further improvement can result from the use of a low-pass filter between the demodulator and the dc indicator such that the signal of selected phase is integrated over an appreciable time interval up to a second or more.

3.1.3 Continuous EMF Measurements

A *standard of emf* may be either an electrochemical system or a Zener-diode-controlled circuit operated under precisely specified conditions. The *Weston standard cell* has a positive electrode of metallic mercury and a negative electrode of cadmium-mercury amalgam (usually about 10% Cd). The *electrolyte* is a saturated solution of cadmium sulfate with an excess of $Cd \cdot SO_4 \cdot {}^8/_3H_2O$ crystals, usually acidified with sulfuric acid (0.04 to 0.08 N). A paste of mercurous sulfate and cadmium sulfate crystals over the mercury electrode is used as a depolarizer. The saturated cell has a substantial temperature coefficient of emf. Vigoureux and Watts of the National Physical Laboratory have given the following formula, applicable to cells with a 10% amalgam:

$$E_t = E_{20} - 39.39 \times 10^{-6}(t - 20) - 0.903 \times 10^{-6}(t - 20)^2 + 0.00660 \times 10^{-6}(t - 20)^3 - 0.000150 \times 10^{-6}(t - 20)^4$$

where t is the temperature in degree Celsius. Since cells are frequently maintained at 28°C, the following equivalent formula is useful:

$$E_t = E_{28} - 52.899 \times 10^{-6}(t - 28) - 0.80265 \times 10^{-6}(t - 28)^2 + 0.001813 \times 10^{-6}(t - 28)^3 - 0.0001497 \times 10^{-6}(t - 28)^4$$

These equations are general and are normally used only to correct cell emfs for small temperature changes, that is, 0.05 K or less. For changes at that level, negligible errors are introduced by making corrections. Standard cells should always be calibrated at their temperature of use (within 0.05 K) if they are to be used at an accuracy of 5 ppm or better.

A group of saturated Weston cells, maintained at a constant temperature in an air bath or a stirred oil bath, is quite generally used as a laboratory reference standard of emf. The bath temperature must be constant within a few thousandths of a degree if the reference emf is to be reliable to a microvolt. It is even more important that temperature gradients in the bath be avoided, since the individual limbs of the cell have very large temperature coefficients (about +315 μV/°C for the positive limb and −379 μV/°C for the negative limb—more than −50 μV/°C for the complete cell—at 28°C). Frequently, two or three groups of cells are used, one as a reference standard which never leaves the laboratory, the others as transport groups which are used for interlaboratory comparisons and for assignment by a standards laboratory.

Precautions in Using Standard Cells

1. The cell should not be exposed to extreme temperatures—below 4°C or above 40°C.

2. Temperature gradients (differences between the cell limbs) should be avoided.

3. Abrupt temperature changes should be avoided—the recovery period after a sudden temperature change may be quite extended; recovery is usually much quicker in an unsaturated than in a saturated cell. Full recovery of saturated cells from a gross temperature change (e.g., from room temperature to a 35°C maintenance temperature) can take up to 3 months. More significantly, some cell emfs have been seen to exhibit a plateau in their response over a 2- to 3-week period within a week or two after the temperature shock is sustained. This plateau can be as much as 5 ppm higher than the final stable value.

4. Current in excess of 100 nA should never be passed through the cell in either direction; actually, one should limit current to 10 nA or less for as short a time as feasible in using the cell as a reference. Cells that have been short-circuited or subjected to excessive charging current drift until chemical equilibrium in the cell is regained over an extended time period—as long as 9 months, depending on the amount of charge involved.

Zener diodes or diode-based devices have replaced chemical cells as voltage references in commercial instruments, such as digital voltmeters and voltage calibrators. Some of these instruments have uncertainties below 10 ppm, instabilities below 5 ppm per month (including drift and random uncertainties), and temperature coefficient of output as low as 2 ppm/°C.

The best devices, as identified in a testing in selection process, are available as solid-state voltage reference or transport standards. Such instruments generally have at least two outputs, one in the range of 1.018 to 1.02 V for use as a standard cell replacement and the other in the range of 6.4 to 10 V, the output voltage of the reference device itself. The lower voltage is usually obtained via a resistive divider.

Other features sometimes include a vernier adjustment for the lower voltage for adjusting to equal the output of a given standard cell and internal batteries for complete isolation. Such devices have performance approaching that of standard cells and can be used in many of the same applications. Some have stabilities (drift rate and random fluctuations) as low as 2 to 3 ppm per year and temperature coefficient of 0.1 ppm/°C.

The current through the reverse-biased junction of a silicon diode remains very small until the bias voltage exceeds a characteristic V_z in magnitude, at which point its resistance becomes abruptly very low so that the voltage across the junction is little affected by the junction current. Since the voltage-current relationship is repeatable, the diode may be used as a standard of voltage as long as its rated power is not exceeded.

However, since V_z is a function of temperature, single junctions are rarely used as voltage references in precise applications. Since a change in temperature shifts the *I-V* curve of a junction, the use of a forward-biased junction in series with Zener diode permits a current level to be found at which changes in Zener voltage from temperature changes are compensated by changes in the voltage drop across the forward-biased junction.

Devices using this principle fall into two categories: the temperature-compensated Zener diode, in which two diodes are in series opposition, and the reference amplifier, in which the Zener diode is in series with the base-emitter junction of an appropriate *npn* silicon transistor. In each case, the two elements may be on the same substrate for temperature uniformity. In some precision devices, the reference element is in a temperature-controlled oven to permit even greater immunity to temperature fluctuations.

Potentiometers are used for the precise measurement of emf in the range below 1.5 V. This is accomplished by opposing to the unknown emf an equal *IR* drop. There are two possibilities: either the current is held constant while the resistance across which the *IR* drop is opposed to the unknown is varied, or current is varied in a fixed resistance to achieve the desired *IR* drop.

Figure 3-5 shows schematically most of the essential features of a general-purpose constant-current instrument. With the standard-cell dial set to read the emf of the reference standard cell, the potentiometer current *I* is adjusted until the *IR* drop across 10 of the coarse-dial steps plus the drop to the set point on the standard-cell dial balances the emf of the reference cell. The correct value of current is indicated by a null reading of the galvanometer in position G_1. This adjustment permits the potentiometer to be read directly in volts. With the galvanometer in position G_2, the unknown emf is balanced by varying the opposing *IR* drop. Resistances used from the *coarse* and *intermediate* dials and the *slide wire* are adjusted until the galvanometer again reads null, and the unknown emf can be read directly from the dial settings. The ratio of the unknown and reference emfs is precisely the ratio as the resistances for the two null adjustments, provided that the current is the same.

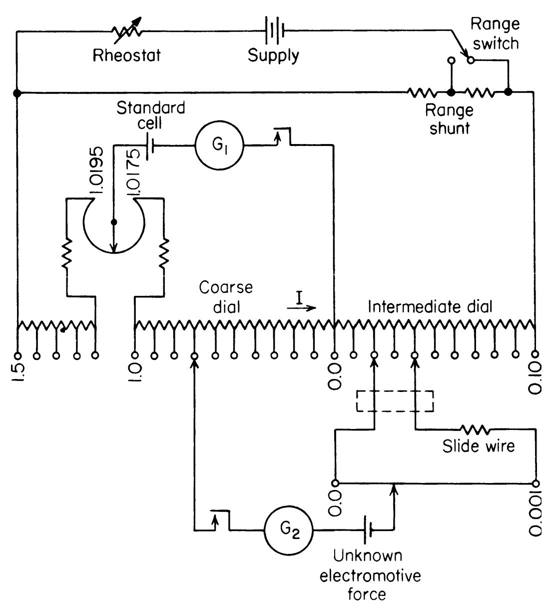


FIGURE 3-5 General purpose constant-current potentiometer.

The switching arrangement is usually such that the galvanometer can be shifted quickly between the G_2 and G_1 positions to check that the current has not drifted from the value at which it was standardized. It will be noted that the contacts of the coarse-dial switch and slide wire are in the galvanometer branch of the circuit. At balance, they carry no current, and their contact resistance does not contribute to the measurement. However, there can be only two noncontributing contact resistances in the network shown; the switch contacts for adjusting the intermediate-dial position do carry current, and their resistance does enter the measurement. Care is taken in construction that the resistances of such current-carrying contacts are low and repeatable, and frequently, as in the example illustrated, the circuit is arranged so that these contributing contacts carry only a fraction of the reference current, and the contribution of their *IR* drop to the measurement is correspondingly reduced.

Another feature of many general-purpose potentiometers, illustrated in the diagram, is the availability of a reduced range. The resistances of the range shunts have such values that at the 0.1 position of the range-selection switch, only a tenth of the reference current goes through the measuring branch of the circuit, and the range of the potentiometer is correspondingly reduced. Frequently, a $\times$ 0.01 range is also available.

In addition to the effect of *IR* drops at contacts in the measuring circuit, accuracy limits are also imposed by thermal emfs generated at circuit junctions. These limiting factors are increasingly important as potentiometer range is reduced. Thus, in low-range or microvolt potentiometers, special care is taken to keep circuit junctions and contact resistances out of the direct measuring circuit as much as possible, to use thermal shielding, and to arrange the circuit and galvanometer keys so that temperature differences will be minimized between junction points that are directly in the measuring circuit. Generally also, in microvolt potentiometers, the galvanometer is connected to the circuit through a special *thermofree* reversing key so that thermal emfs in the galvanometer can be eliminated from the measurement—the balance point being that which produces zero change in galvanometer deflections on reversal.

An example of the constant-resistance potentiometer is shown in the simplified diagram in Fig. 3-6. It consists basically of a constant-current source, a resistive divider *D* (used in the current-divider mode), and a fixed resistor *R* in which the current (and the *IR* drop) are determined by the setting of the divider. The output of the current source is adjusted by equating the emf of a standard cell to an equal *IR* drop as shown by the dashed line. This design lends itself to multirange operation by using tap points on the resistor *R*. Its accuracy depends on the uniformity of the divider, the location of the tap points on *R*, and the stability of the current source.

Another type of constant-resistance potentiometer, operating from a current comparator which senses and corrects for inequality of ampere-turns in two windings threading a magnetic core, is shown in Fig. 3-7. Two matched toroidal cores wound with an identical number of turns are excited by a fixed-frequency oscillator. The fluxes induced in the cores are equal and oppositely directed, so they cancel with respect to a winding that encloses both. In the absence of additional magnetomotive force (mmf), the detector winding enclosing both cores receives no signal.

If, in another winding *A* enclosing both cores, we inject a direct current, its mmf reinforces the flux in one core and opposes the other. The net flux in the detector winding induces a voltage in it. This signal is used to control current in another winding *B* which also threads both cores. When the mmf of *B* is equal to and opposite that of *A*, the detector signal is zero and the ampere-turns of *A* and *B* are equal. Thus, a constant current in an adjustable number of turns is matched to a variable current in a fixed number of turns, and the voltage drop I_BR is used to oppose the emf to be measured.

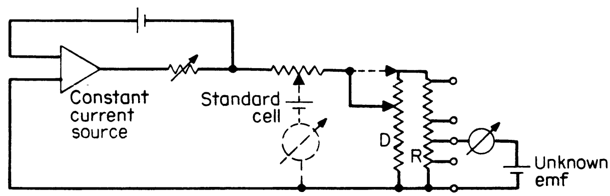


FIGURE 3-6 Constant-resistance potentiometer.

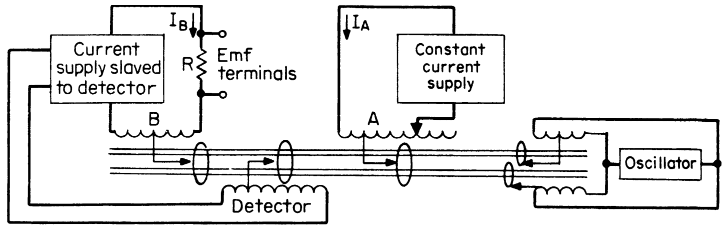


FIGURE 3-7 Current-comparator potentiometer.

The system is made direct-reading in voltage units (in terms of the turns ratio *B*/*A*) by adjusting the constant-current source with the aid of a standard-cell circuit (not shown in the figure). This type of potentiometer has an advantage over those whose continuing accuracy depends on the stability of a resistance ratio; the ratio here is the turns ratio of windings on a common core, dependent solely on conductor position and hence not subject to drift with time.

Decade voltage dividers generally use the Kelvin-Varley circuit arrangement shown in Fig. 3-8. It will be seen that two elements of the first decade are shunted by the entire second decade, whose total resistance equals the combined resistance of the shunted steps of decade I. The two sliders of decade I are mechanically coupled and move together, keeping the shunted resistance constant regardless of switch position. Thus, the current divides equally between decade II and the shunted elements of decade I, and the voltage drop in decade II equals the drop in one unshunted step of decade I. The effect of contact resistance at the switch points is somewhat diminished because of the division of current. The Kelvin-Varley principle is used in succeeding decades except the final one, which has only a single switch contact. Such voltage dividers may have as many as eight decades and have ratio accuracies approaching 1 part in 10^6 of input.

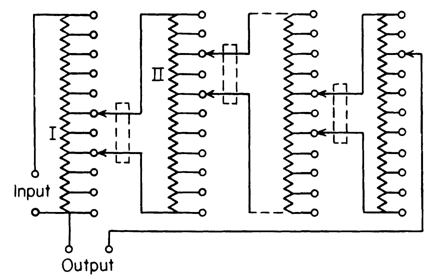


FIGURE 3-8 Decade voltage divider.

Spark gaps provide a means of measuring high voltages. The maximum gap which a given voltage will break down depends on air density, gap geometry, crest value of the voltage, and other factors (see Sec. 27). Sphere gaps constitute a recognized means for measuring crest values of alternating voltages and of impulse voltages. IEEE Standard 4 has tables of sparkover voltages for spheres ranging from 6.25 to 200 cm in diameter and for voltages from 17 to 2500 kV. Sphere gap voltage tables are also available in ANSI Standard 68.1 and in IEC Publication 52.

3.1.4 Continuous Current Measurements

Absolute current measurement relates the value of the current unit—the ampere—to the prototype mechanical units of length, mass, and time—the meter, the kilogram, and the second—through force measurements in an instrument called a *current balance*. Such instruments are to be found generally only in national standards laboratories, which have the responsibility of establishing and maintaining the electrical units. In a current balance, the force between fixed and movable coils is opposed by the gravitational force on a known mass, the balance equation being $I^2(\partial M/\partial X) = mg$. The construction of the coil system is such that the rate of change with displacement of mutual inductance between fixed and moving coils can be computed from measured coil dimensions. Absolute current determinations are used to assign the emf of reference standard cells. A 1-Ω resistance standard is connected in series with the fixed- and moving-coil system, and its drop is compared with the emf of a cell during the force measurement. Thus, the National Reference Standard of voltage is derived from absolute ampere and ohm determinations.

The *potentiometer method* of measuring continuous currents is commonly used where a value must be more accurate than can be obtained from the reading of an indicating instrument. The current to be measured is passed through a four-terminal resistor (shunt) of known value, and the voltage developed between its potential terminals is measured with a potentiometer. If the current is small so that there is no significant temperature rise in the shunt, the measurement accuracy can be 0.01% or better. In general, the accuracy of potentiometer measurements of continuous currents is limited by how well the shunt resistance is known under operating conditions.

Measurement of very small continuous currents, down to 10^{-17} A, have been accomplished by means of *electrometer* tubes—vacuum tubes designed so that the grid has practically no leakage current either over its insulating supports or to the cathode. The current to be measured flows through a very high resistance (up to 10^{12} Ω), and the voltage drop is impressed on the grid of an electrometer tube. The plate current is observed and the voltage drop is duplicated by producing the plate current with a known adjustable voltage. The current can then be calculated from the voltage and resistance.

3.1.5 Analog Instruments

Analog instruments are electromechanical devices in which an electrical quantity is measured by conversion to a mechanical motion. Such instruments can be classified according to the principle on which the instrument operates. The usual types are permanent-magnet moving-coil, moving-iron, dynamometer, and electrostatic. Another grouping is on the basis of use: panel, switchboard, portable, and laboratory-standard. Accuracy also can be the basis of classification. Details concerning performance and other specifications are to be found in ANSI Standard C39.1, Requirements for Electrical Analog Indicating Instruments.

Permanent-magnet moving-coil instruments are the most common type in general use. The operating mechanism consists of a coil of fine wire suspended in such a manner that it can rotate in an annular gap which has a radial magnetic field. The torque, generated by the current in the moving coil reacting to the magnetic field of the gap, is opposed by some form of spring restraint. The restraint may be a helical spring, in which case the coil is supported by a pivot and jewel, or both the support and the angular restraint is by means of a taut-band suspension.

The position which the coil assumes when the torque and spring restraint are balanced is indicated by either a pointer or a light beam on a scale. The scale is calibrated in units suitable to the application: volts, milliamperes, etc. To the extent that the magnetic field is uniform, the spring restraint linear, and the coil positioning symmetrical, the deflection will be linearly proportional to the ampere-turns in the coil.

Because the field of the permanent magnet is unidirectional, reversal of the coil current will reverse the torque so that the instrument will deflect only with direct current in the moving coil. Scales are usually provided with the zero-current position at the left to allow a full-range deflection. However, where measurement is required with either polarity, a zero center scale position is used. The coil is limited in its ability to carry current to 50 or 100 mA.

Rectifiers and thermoelements are used with permanent-magnet moving-coil instruments to provide ac operation. The addition of a rectifier circuit, usually in the form of a bridge, gives an instrument in which the deflection is in terms of the average value of the voltage or current. It is customary to label the scale in terms of 1.11 times the average; this is the correct waveform factor to read the rms value of a sine wave. If the rectifier instrument is used to measure severely nonsinusoidal waveforms, large errors will result. The high sensitivity that can be obtained with the rectifier type of instrument and its reasonable cost make it widely used.

To provide a true rms reading with the permanent-magnet moving-coil instrument, a thermoelement is the usual converter. The current to be measured is fed through a resistance of such value that it will heat appreciably. A thermocouple is placed in intimate thermal contact with the heater resistance, and the output of the couple is used to energize a permanent-magnet moving-coil instrument. The instrument deflection of such a combination is proportional to the square of the current; using a square-root factor in drawing the scale allows it to be read in terms of the rms value of the current. For high-sensitivity use, the thermoelement is placed in an evacuated bulb to eliminate convection heat loss.

The prime advantage of the thermoelement instrument is the high frequency at which it will operate and the rms indication. The upper frequency limit is determined by the skin effect in the heater. Instruments have been built with response to several hundred megahertz. There is one very important limitation to these instruments. The heater must operate at a temperature of 100°C or more to provide adequate current to the movement. Overrange of the current will cause heater temperature to increase as the square of the current. It is possible to burn out the heater with relatively small overloads.

Moving-iron instruments are widely used at power frequencies. The radial-vane moving-iron type operates by current in the coil which surrounds two magnetic vanes, one fixed and one that can rotate in such a manner as to increase the spacing between them. Current in the coil causes the vanes to be similarly magnetized and so to repel each other. The torque produced by the moving vane is proportional to the square of the current and is independent of its polarity.

Figure 3-9 shows two ways in which a wattmeter may be connected to measure power in a load. With the moving coil connected at *A*, the instrument will read high by the amount of power used by the moving-coil circuit. If connection is made at *B*, the wattmeter will read high by the power dissipated in the field coils. When using sensitive, low-range meters, it is necessary to correct for this error. Commercial instruments are available for ranges from a fraction of a watt to several hundred watts self-contained. Range extensions are obtained with current and voltage transformers. In specifying wattmeters, it is necessary to state the current and voltage ranges as well as the watt range.

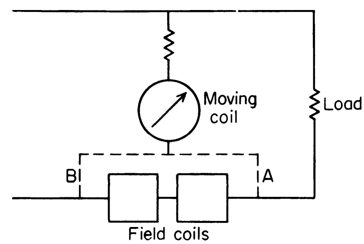


FIGURE 3-9 Alternative wattmeter connections.

Electrostatic voltmeters are actually voltage-operated in contrast to all the other types of analog instruments, which are current-operated. In an electrostatic voltmeter, fixed and movable vanes are so arranged that a voltage between them causes attraction to rotate the movable vane. The torque is proportional to the energy stored in the capacitance, and thus to the voltage squared, permitting rms indication.

Electrostatic instruments are used for voltage measurements where the current drain of other types of instrument cannot be tolerated. Input resistance (due to insulation leakage) amounts to $10^{13}\ \Omega$ approximately for a range of 100 V (the lowest commercially available) to $3 \times 10^{15}\ \Omega$ for 100,000-V instruments (the highest commonly available). Capacitance ranges from about 300 pF for the lower ranges to 10 pF for the highest. Multirange instruments in the lower ranges (100 to 5000 V) are frequently made with capacitive dividers which make them inoperable on direct voltage, since the series capacitor blocks out dc. Other multirange instruments use a mechanical movement of the fixed electrode to change ranges. These can be used on dc or ac, as can all single-range voltmeters.

Electronic voltmeters vary widely in performance characteristics and frequency range covered, depending on the circuitry used. A common type uses an initial diode to charge a capacitor. This may be followed by a stabilized amplifier with a microammeter as indicator. Range may be selected by appropriate cathode resistors in the amplifier section. Such instruments normally have very high input impedance (a few picofarads), respond to peak voltage, and are suitable for use to very high frequencies (100 MHz or more). While the response is to *peak* voltage, the scale of the indicating element may be marked in terms of rms for a sine-wave input, that is, 0.707 × *peak* voltage. Thus, for a nonsinusoidal input, the scale (read as rms volts) may include a serious waveform error, but if the scale reading is multiplied by 1.41, the result is the value of the *peak* voltage.

An alternative network, used in some electronic voltmeters, is an attenuator for range selection, followed by an amplifier and finally a rectifier and microammeter. This system has substantially lower input impedance, and limits of frequency range are fixed by the characteristics of the amplifier. The response in this arrangement may be to *average* value of the input signal, but again, the scale marking may be in terms of rms value for a sine wave. In this case also, the waveform error for nonsinusoidal input must be borne in mind, but if the scale reading is divided by 1.11, the *average* value is obtained. Within these limitations, accuracy may be as good as 1% of full-scale indication in sometypes of electronic voltmeter, although in many cases a 2 to 5% accuracy may be anticipated.

3.1.6 DC to AC Transfer

General transfer capability is essential to the measurement of voltage, current, power, and energy. The standard cell, the unit of voltage which it preserves, and the unit of current derived from it in combination with a standard of resistance are applicable only to the measurement of dc quantities, while the problems of measurement in the power and communications fields involve alternating voltages and currents. It is only by means of transfer devices that one can assign the values of ac quantities or calibrate ac instruments in terms of the basic dc reference standards. In most instances, the rms value of a voltage or current is required, since the transformation of electrical energy to other forms involves the square of voltages or currents, and the transfer from direct to alternating quantities is made with devices that respond to the square of current or voltage. Three general types of transfer instruments are capable of high-accuracy rms measurements: (1) electrodynamic instruments—which depend on the force between current-carrying conductors; (2) electrothermic instruments—which depend on the heating effect of current; and (3) electrostatic instruments—which depend on the force between electrodes at different potentials. While two of these depend on current and the third on voltage, the use of series and shunt resistors makes all three types available for current or voltage transfer. Traditional American practice has been to use electrodynamic instruments for current and voltage transfer as well as power transfer from direct to alternating current, but recent developments in thermoelements have improved their transfer characteristics until they are now the preferred means for *current* and *voltage* transfer, although the electrodynamic wattmeter is still the instrument of choice for *power* transfer up to 1 kHz.

Electrothermic transfer standards for current and voltage use a thermoelement consisting of a heater and a thermocouple. In its usual form, the heater is a short, straight wire suspended by two supporting lead-in wires in an evacuated glass bulb. One junction of a thermocouple is fastened to its midpoint and is electrically insulated from it with a small bead. The thermal emf—5 to 10 mV at rated current in a conventional element—is a measure of heater current. Multijunction thermoelements having a number of couples in series along the heater also have been used in transfer measurements. Typical output is 100 mV for an input power of 30 mW.

3.1.7 Digital Instruments

Digital voltmeters (DVMs), displaying the measured voltage as a set of numerals, are analog-to-digital converters in which an unknown dc voltage is compared with a stable reference voltage. Internal fixed dividers or amplifiers extend the voltage ranges. For ac measurements, dc DVMs are preceded by ac-to-dc converters. DVMs are widely used as laboratory, portable, and panel instruments because of their convenience, accuracy, and speed. Automatic range changing and polarity indication, freedom from reading errors, and the availability of outputs for data acquisition or control are added advantages. Integrated circuits and modern techniques have greatly increased their reliability and reduced their cost. Full-scale accuracies range from about 0.5% for three-digit panel instruments to 1 ppm for eight-digit laboratory dc voltmeters and 0.016% for ac voltmeters.

Successive-approximation DVMs are automatically operated dc potentiometers. These may be based on resistive voltage or current divider techniques or on dc current comparators. A comparator in a series of steps adjusts a discrete fraction of the reference voltage (by current or voltage division in a resistance network) until it equals the unknown. Various "logic schemes" have been used to accomplish this, and the stepping relays of earlier models have been replaced by electronic or reed switches. Filters reduce input noise (which could prevent a final display) but generally increase the response time. Accuracy depends chiefly on the reference voltage and the ratios of the resistance network.

Voltage-to-frequency-converter (V/f) DVMs generate a ramp voltage at a rate proportional to the input until it equals a fixed voltage, returns the ramp to the starting point, and repeats. The number of pulses (ramps) generated in a fixed time is proportional to the input and is counted and displayed. Since it integrates over the counting time, a *V/f* DVM has excellent input-noise rejection. The ramp is usually generated by an operational integrator (a high-gain operational amplifier with a capacitor in the feedback loop so that its output is proportional to the integral of the input voltage). The capacitor is discharged each time by a pulse of constant and opposite charge, and the time interval of the counter is chosen so that the number of pulses makes the DVM direct-reading. Accuracy depends on the integrator and on the charge of the pulse generator, which contains the reference voltage.

Dual-slope DVMs generate a voltage ramp at a rate proportional to the input voltage V_i for a fixed time t_1. The ramp input is then switched to a reference voltage V_r of the opposite polarity for a time t_2 until the starting level is reached. Pulses with a fixed frequency f are accumulated in a counter, with N_1 counts during t_1. The counter resets to zero and accumulates N_2 counts during t_2. Thus, $t_1 = N_1 f$ and $t_2 = N_2 f$.

If the slope of the linear ramp is $m = kV$, the ramp voltage is $V_o = mt = kVt$. Thus $V_i t_1 = V_r t_2$, so $V_i = V_r N_2/N_1$. The time t_1 is controlled by the counter to make N_2 direct-reading in appropriate units. In principle, the accuracy is not dependent on the constants of the ramp generator or the frequency of the pulses. A single operational integrator, switched to either input or reference voltage, generates the ramps. Since there are few critical components, integrated circuits are feasible, leading to simplicity and reliability as well as high accuracy. Because this is an integrating DVM, noise rejection is excellent.

In pulse-width conversion meters, an integrating circuit and matched comparators are used to produce trains of positive and negative pulses whose relative widths are a linear function of any dc input. The difference in positive and negative pulse widths can be measured using counting techniques, and very high resolution and accuracy (up to 1 ppm, relative to an internal voltage reference) can be achieved by integrating the counting over a suitable time period.

Average ac-to-dc converters contain an operational rectifier (an operational amplifier with a rectifier in the feedback circuit), followed by a filter, to obtain the rectified average value of the ac voltage. The operational amplifier greatly reduces errors of nonlinearity and forward voltage drop of the rectifier. For convenience, the output voltage is scaled so that the dc DVM connected to it indicates the rms value of a sine wave. Large errors can result for other waveforms, up to h/n%, with h% of the nth harmonic in the wave, if n is an odd number. For example, with 3% of third harmonic, the error can be as much as 1%, depending on the phase of the harmonic.

Electronic multipliers and other forms of rms-responding ac-to-dc converters eliminate this waveform error but are generally more complex and expensive. In one version, the feedback rms circuit shown in Fig. 3-10, the two inputs of the multiplier M_1 are connected together so that the instantaneous output of M is v_i^2/V_o. The operational filter F(RC circuit and operational amplifier) makes $V_o = V_i^2/V_o$, where V_i^2 is the square of the rms value. Thus, $V_o = V_i$. The conversion accuracy approaches 0.1% up to 20 kHz in transconductance or logarithmic multipliers, without requiring a wide dynamic range in the instrument, because of the internal feedback. A series of diodes, biased to conduct at different voltage levels, can provide an excellent approximation to a square-law function in a feedback circuit like that of Fig. 3-10.

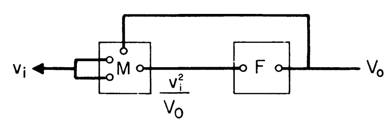


FIGURE 3-10 Electronic rms ac-to-dc converter.

Specifications for DVMs should follow the recommendations of ANSI Standard C39.7, Requirements for Digital Voltmeters. Accuracy should be stated as the overall limit of error for a specified range of operating conditions. It should be in percent of reading plus percent of full scale and may be different for different frequency and voltage ranges. Accuracy at a narrow range of reference conditions is also often specified for laboratory use. The input configuration (two-terminal, three-terminal unguarded, three- or four-terminal guarded) is important. Number of digits and "overrange" also should be stated.

Errors and Precautions. Because of the sensitivity of DVMs, a number of precautions should be taken to avoid in-circuit errors from ground loops, input noise, etc. The high input impedance of most types makes input loading errors negligible, but this should always be checked. On dc millivolt ranges, unwanted thermal emfs should be checked as well as the normal-mode rejection of ac line-frequency voltage across the input terminals. Two-terminal DVMs (chassis connected to one input as well as to line ground) may measure unwanted voltages from ground currents in the common line.

Errors are greatly reduced in three-terminal DVMs (chassis connected to line ground only) and are generally negligible with guarded four-terminal DVMs (separate guard chassis surrounding the measuring circuit). Such DVMs have very high common-mode rejection. Some types of DVMs introduce small voltage spikes or currents to the measuring circuit, often from internal switching transients, which may cause errors in low-level circuits.

Digital multimeters are DVMs with added circuitry to measure quantities such as dc voltage ratio, dc and ac current, and resistance. Voltage ratio is measured by replacing the reference voltage with one of the unknowns. For current, the voltage across an internal resistor carrying the current is measured by the DVM. For resistance, a fixed reference current is generated and applied to the unknown resistor. The voltage across the resistor is measured by the DVM. Several ranges are provided in each case.

3.1.8 Instrument Transformers

The material that follows is a brief summary of information on instrument transformers as measurement elements. For more extensive information, consult American National Standard C57.13, Requirement for Instrument Transformers; American National Standards Institute; American National Standard C12, Code for Electricity Metering; *Electrical Meterman's Handbook*, Edison Electric Institute; manufacturer's literature; and textbooks on electrical measurements.

AC range extension beyond the reasonable capability of indicating instruments is accomplished with instrument transformers, since the use of heavy-current shunts and high-voltage multipliers would be prohibitive both in cost and power consumption. Instrument transformers are also used to isolate instruments from power lines and to permit instrument circuits to be grounded.

The current circuits of instruments and meters normally have very low impedance, and current transformers must be designed for operation into such a low-impedance secondary burden. The insulation from the primary to secondary of the transformer must be adequate to withstand line-to-ground voltage, since the connected instruments are usually at ground potential. Normal design is for operation with a rated secondary current of 5 A, and the input current may range upward to many thousand amperes. The potential circuits of instruments are of high impedance, and voltage transformers are designed for operation into a high-impedance secondary burden. In the usual design, the rated secondary voltage is 120 V, and instrument transformers have been built for rated primary voltages up to 765 kV.

With the development of higher transmission-line voltages (350 to 765 kV) and intersystem ties at these levels, the coupling-capacitor voltage transformer (CCVT) has come into use for metering purposes to replace the conventional voltage transformer, which, at these voltages, is bulkier and more costly. The metering CCVT, shown in Fig. 3-11, consists of a modular capacitive divider which reduces the line voltage V_1 to a voltage V_2 (10–20 kV), with a series-resonant inductor to tune out the high impedance and make available energy transfer across the divider to operate the voltage transformer which further reduces the voltage to V_M, the metering level. Required metering accuracy may be 0.3% or better.

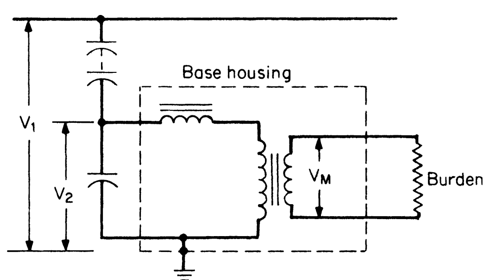


FIGURE 3-11 CCTV metering arrangement.

Instrument transformers are broadly classified in two general types: (1) dry type, having molded insulation (sometimes only varnish-impregnated paper or cloth) usually intended for indoor installation, although large numbers of modern transformers have molded insulation suitable for outdoor operation on circuits up to 15 kV to ground; and (2) liquid-filled types in steel tanks with high-voltage primary terminals, intended for installation on circuits above 15 kV. They are further classified according to accuracy: (1) metering transformers having highest accuracy, usually at relatively low burdens; and (2) relaying and control transformers which in general have higher burden capacity and lower accuracy, particularly at heavy overloads. This accuracy classification is not rigid, since many transformers, often in larger sizes and higher voltage ratings, are suitable for both metering and control purposes.

Another classification differentiates between single and multiple ratios. Multiple primary windings, sometimes arranged for series-parallel connection, tapped primary windings, or tapped secondary windings, are employed to provide multiple ratios in a single piece of equipment. Current transformers are further classified according to their mechanical structure: (1) wound primary, having more than one turn through the core window; (2) through type, wherein the circuit conductor (cable or busbar) is passed through the window; (3) bar type, having a bar, rod, or tube mounted in the window; and (4) bushing type, that is, through type intended for mounting on the insulating bushing of a power transformer or circuit breaker.

Current transformers, whose primary winding is series connected in the line, serve the double purposes of (1) convenient measurement of large currents and (2) insulation of instruments, meters, and relays from high-voltage circuits. Such a transformer has a high-permeability core of relatively small cross section operated normally at a very low flux density. The secondary winding is usually in excess of 100 turns (except for certain small low-burden through-type current transformers used for metering, where the secondary turns may be as low as 40), and the primary is of few turns and may even be a single turn or a section of a bus bar threading the core. The nominal current ratio of such a transformer is the inverse of the turns ratio, but for accurate current measurement, the actual ratio must be determined under loading corresponding to use conditions. For accurate power and energy measurement, the phase angle between the secondary and reversed primary phasor also must be known for the use condition. Insulation of primary from secondary and core must be sufficient to withstand, with a reasonable safety factor, the voltage to ground of the circuit into which it is connected; secondary insulation is much less, since the connected instrument burden is at ground potential or nearly so.

The overload capacity of station-type current transformers and the mechanical strength of the winding and core structure must be high to withstand possible short circuits on the line. Various compensation schemes are used in many transformers to retain ratio accuracy up to several times rated current. The secondary circuit—the current elements of connected instruments or relays—*must never be opened* while the transformer is excited by primary current, because high voltages are induced which may be hazardous to insulation and to personnel and because the accuracy of the transformer may be adversely affected.

Voltage transformers (potential transformers) are connected between the lines whose potential difference is to be determined and are used to step the voltage down (usually to 120 V) and to supply the voltage circuits of the connected instrument burden. Their basic construction is similar to that of a power transformer operating at the same input voltage, except that they are designed for optimal performance with the high-impedance secondary loads of the connected instruments. The core is operated at high flux density, and the insulation must be appropriate to the line-to-ground voltage.

Standard burdens and standard accuracy requirements for instrument transformers are given in American National Standard C57.13 (see Sec. 28).

Accuracy. Most well-designed instrument transformers (provided they have not been damaged or incorrectly used) have sufficient accuracy for metering purposes. See Sec. 10 for typical accuracy curves. Where higher accuracy is required, see Appendix D of ANSI C12, The Code for Electricity Metering.

Another comparison method uses a “standard” transformer of the same nominal rating as the one being tested. Accuracies of 0.01% are attainable. Commercial test sets are available for this work and are widely used in laboratory and field tests. Commercial test sets based on the current-comparator method and capable of 0.001% accuracy are also available. For further details, see ANSI Standard C57.13.

3.1.9 Power Measurement

Electronic wattmeters of 0.1% or better accuracy may be based on a pulse-area principle. Voltages proportional to the applied voltage and to the current (derived from resistors or transformers) govern the height and width of a rectangular pulse so that the area is proportional to the instantaneous power. This is repeated many times during a cycle, and its average represents active power. Average power also can be measured by a system which samples instantaneous voltage and current repeatedly, at predetermined intervals within a cycle. The sampled signals are digitized, and the result is computed by numerical integration. The response of such a system has been found to agree with that of a standard electrodynamic wattmeter within 0.02% from dc to 1 kHz. Depending on sampling speed, measurements can be made to higher frequencies with somewhat reduced accuracy. In the digital instrument, the multiplication involves discrete numbers and thus has no experimental error except for rounding. Such an arrangement is well-adapted to the measurement of power in situations where current or voltage waveforms are badly distorted.

In the thermal wattmeter, where the arrangement is such that if one current v is proportional to instantaneous load voltage and another i is proportional to load current, their sum is applied to one thermal converter and their difference to another. Assuming identical quadratic response of the converters, their differential output may be represented as

$$V_{dc} = \frac{k}{T}\int_0^T [(v + i)^2 - (v - i)^2]dt = \frac{4k}{T}\int_0^T vi\,dt$$

which is by definition average power. Multijunction thermal converters with outputs connected differentially are used for the ac-dc transfer of power, with ac and dc current and voltage signals applied simultaneously to both heaters. DC feedback to current input speeds response and maintains thermal balance between heaters, and the output meter becomes a null indicator. This mode of operation can eliminate the requirement for exact quadrature response, and the matching requirement is also eliminated by interchange of the heaters. The Cox and Kusters instrument was designed for operation from 50 to 1000 Hz with ac-dc transfer errors within 30 ppm, and it may be used up to 20 kHz with reduced accuracy. This instrument also is capable of precision measurement with very distorted waveforms.

Laboratory-standard wattmeters use an electrodynamic mechanism and are in the 0.1% accuracy class for dc and for ac up to 133 Hz. This accuracy can be maintained up to 1 kHz or more. Such instruments are shielded from the effects of external magnetic fields by enclosing the coil system in a laminated iron cylinder. Instruments having current ranges to 10 A and voltage ranges to 300 V are generally self-contained. Higher ranges are realized with the aid of precision instrument transformers.

Portable wattmeters are generally of the electrodynamic type. The current element consists of two fixed coils connected in series with the load to be measured. The *voltage element* is a moving coil supported on jewel bearings or suspended by taut bands between the fixed field coils. The moving coil is connected in series with a relatively large noninductive resistor across the load circuit. The coils are mounted in a laminated iron shield to minimize coupling with external magnetic fields. Switchboard wattmeters have the same coil structure but are of broader accuracy class and do not have the temperature compensation, knife-edge pointers, and antiparallax mirrors required for the better-class portable instruments.

Correction for wattmeter power consumption may be important when the power measured is small. When the wattmeter is connected directly to the circuit (without the interposition of instrument transformers), the instrument reading will include the power consumed in the element connected next to the load being measured. If the instrument loss cannot be neglected, it is better to connect the voltage circuit next to the load and include its power consumption rather than that of the current circuit, since it is generally more nearly constant and is more easily calculated. In some low-range wattmeters, designed for use at low-power factors, the loss in the voltage circuit is automatically compensated by carrying the current of the voltage circuit through compensating coils wound over the field coils of the current circuit. In this case, the voltage circuit must be connected next to the load to obtain compensation.

The *inductance error* of a wattmeter may be important at low-power factor. At power factors near unity, the noninductive series resistance in the voltage circuit is large enough to make the effect of the moving-coil inductance negligible at power frequencies, but with low power factor, the phase angle of the voltage circuit may have to be considered. This may be computed as $\alpha = 21.6fL/R$, where α is the phase angle in minutes, f is the frequency in hertz, L is the moving-coil inductance in millihenrys, and R is the total resistance of the voltage circuit in ohms.

A *3-phase 3-wire circuit* requires two wattmeters connected as shown in Fig. 3-12; total power is the algebraic sum of the two readings under all conditions of load and power factor. If the load is balanced, at unity power factor each instrument will read half the load; at 50% power factor one instrument reads all the load and the other reading is zero; at less than 50% power factor one reading will be negative.

Three-phase 4-wire circuits require three wattmeters as shown in Fig. 3-13. Total power is the algebraic sum of the three readings under all conditions of load and power factor. A 3-phase Y system with

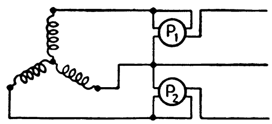


FIGURE 3-12 Power in 3-phase, 3-wire circuit, two wattmeters.

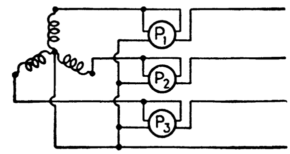


FIGURE 3-13 Power in 3-phase, 4-wire circuit, three wattmeters.

a grounded neutral is the equivalent of a 4-wire system and requires the use of three wattmeters. If the load is balanced, one wattmeter can be used with its current coil in series with one conductor and the voltage circuit connected between that conductor and the neutral. Total power is three times the wattmeter reading in this instance.

Reactive power (reactive voltamperes, or vars) is measured by a wattmeter with its current coils in series with the circuit and the current in its voltage element in quadrature with the circuit voltage.

Corrections for instrument transformers are of two kinds. *Ratio* errors, resulting from deviations of the actual ratio from its nominal, may be obtained from a calibration curve showing true ratio at the instrument burden imposed on the transformer and for the current or voltage of the measurement. The effect of *phase-angle* changes introduced by instrument transformers is modification in the angle between the current in the field coils and the moving coil of the wattmeter; the resulting error depends on the power factor of the circuit and may be positive or negative depending on phase relations, as shown in the table below. If $\cos\theta$ is the *true* power factor in the circuit and $\cos\theta_2$ is the *apparent* power factor (i.e., as determined from the wattmeter reading and the secondary voltamperes), and if K_c and K_v are the true ratios of the current and voltage transformers, respectively, then

$$\text{Main-circuit watts} = K_c K_v \frac{\cos\theta}{\cos\theta_2} \times \text{wattmeter watts}$$

The line power factor $\cos\theta = \cos(\theta_2 \pm \alpha \pm \beta \pm \gamma)$, where θ_2 is the phase angle of the secondary circuit, α is the angle of the wattmeter's voltage circuit, β is the phase angle of the current transformer, and γ is the phase angle of the voltage transformer. These angles—α, β, and γ—are given positive signs when they act to decrease and negative when they act to increase the phase angle between instrument current and voltage with respect to that of the circuit. This is so because a decreased phase angle gives too large a reading and requires a negative correction (and vice versa), as shown in the following table of signs.

Dielectric loss, which occurs in cables and insulating bushings used at high voltages, represents an undesirable absorption of available energy and, more important, a restriction on the capacity of cables and insulating structures used in high-voltage power transmission. The problem of measuring the power consumed in these insulators is quite special, since their power factor is extremely low and the usual wattmeter techniques of power measurement are not applicable. While many methods have been devised over the past half century for the measurement of such losses, the Schering bridge is almost universally the method of choice at the present time. Figure 3-14 shows the basic circuit of the bridge, as described by Schering and Semm in 1920. The balance equations are $C_x = C_s R_1/R_2$ and $\tan\delta_x = \omega R_1 P$, where C_x is the cable or bushing whose losses are to be determined, C_s is a loss-free high-voltage air-dielectric capacitor, R_1 and R_2 are noninductive resistors, and P is an adjustable low-voltage capacitor having negligible loss.

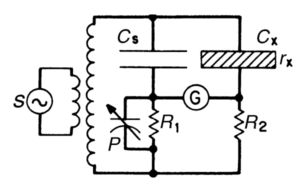


FIGURE 3-14 Schering and Semm's bridge for measuring dielectric loss.

3.1.10 Power-Factor Measurement

The *power factor of a single-phase circuit* is the ratio of the true power in watts, as measured with a wattmeter, to the apparent power in voltamperes, obtained as the product of the voltage and current.

Line power factor	Sign to be used for phase angle					
	α wattmeter		β current transf.		γ voltage transf.	
	Lead*	Lag	Lead	Lag	Lead	Lag
Lead	+	−	−	+	+	−
Lag	−	+	+	−	−	+

*In general, α will be leading only when the inductance of the potential coil has been overcompensated with capacitance.

When the waveform is sinusoidal (and only then), the power factor is also equal to the cosine of the phase angle.

The *power factor of a polyphase circuit* which is balanced is the same as that of the individual phases. When the phases are not balanced, the true power factor is indeterminate. In the wattmeter-voltmeter-ammeter method, the power factor for a balanced 2-phase 3-wire circuit is $P/(\sqrt{2}EI)$, where P is total power in watts, E is voltage between outside conductors, and I is current in an outside conductor; for a balanced 3-phase 3-wire circuit, the power factor is $P/(\sqrt{3}EI)$, where P is total watts, E is volts between conductors, and I is amperes in a conductor. In the two-wattmeter method, the power factor of a 2-phase 3-wire circuit is obtained from the relation $W_2/W_1 = \tan\theta$, where W_1 is the reading of a wattmeter connected in one phase as in a single-phase circuit and W_2 is the reading of a wattmeter connected with its current coil in series with that of W_1 and its voltage coil across the second phase. At unity power factor, $W_2 = 0$; at 0.707 power factor, $W_2 = W_1$; at lower power factors, $W_2 > W_1$. In a 3-phase 3-wire circuit, power factor can be calculated from the reading of two wattmeters connected in the standard way for measuring power, by using the relation

$$\tan\theta = \frac{\sqrt{3}(W_1 - W_2)}{W_1 + W_2}$$

where W_1 is the larger reading (always positive) and W_2 the smaller.

Power-factor meters, which indicate the power factor of a circuit directly, are made both as portable and as switchboard types. The mechanism of a single-phase electrodynamic meter resembles that of a wattmeter except that the moving system has two coils M, M'. One coil, M, is connected across the line in series with a resistor, whereas M' is connected in series with an inductance. Their currents will be nearly in quadrature. At unity power factor, the reaction with the current-coil field results in maximum torque on M, moving the indicator to the 100 mark on the scale, where torque on M is zero. At zero power factor, M' exerts all the torque and causes the moving system to take a position where the plane of M' is parallel to that of the field coils, and the scale indication is zero. At intermediate power factors, both M and M' contribute torque, and the indication is at an intermediate scale position. In a 2-phase meter, the inductance is not required, coil M being connected through a resistance to one phase, while M' with a resistance is connected to the other phase; the current coil may go in the middle conductor of a 3-wire system. Readings are correct only on a balanced load. In one form of polyphase meter, for balanced circuits, there are three coils in the moving system, connected one across each phase. The moving system takes a position where the resultant of the three torques is minimum, and this depends on power factor. In another form, three stationary coils produce a field which reacts on a moving voltage coil. When the load is unbalanced, neither form is correct.

3.1.11 Energy Measurements

The subject of metering electric power and energy is extensively covered in the American National Standard C21, Code for Electricity Metering, American National Standards Institute. It covers definitions, circuit theory, performance standards for new meters, test methods, and installation standards for watthour meters, demand meters, pulse recorders, instrument transformers, and auxiliary devices. Further detailed information may be found in the *Handbook for Electric Metermen,* Edison Electric Institute.

The practical unit of electrical energy is the watthour, which is the energy expended in 1 h when the power (or rate of expenditure) is 1 W. Energy is measured in watthours (or kilowatthours) by means of a *watthour meter*. A watthour meter is a motor mechanism in which a rotor element revolves at a speed proportional to power flow and drives a registering device on which energy consumption is integrated. Meters for continuous current are usually of the mercury-motor type, whereas those for alternating current utilize the principle of the induction motor.

Polyphase Meter Connections. Obviously, it is extremely important that the various circuits of a polyphase meter be properly connected. If, for example, the current-coil connections are interchanged and the line power factor is 50%, the meter will run at the normal 100% power-factor speed, thus giving an error of 100%.

A test for correct connections is as follows: If the line power factor is over 50%, rotation will always be forward when the potential or the current circuit of either element is disconnected, but in one case the speed will be less than in the other. If the power factor is less than 50%, the rotation in one case will be backward.

When it is not known whether the power factor is less or greater than 50%, this may be determined by disconnecting one element and noting the speed produced by the remaining element. Then change the voltage connection of the remaining element from the middle wire to the other outside wire and again note the speed. If the power factor is over 50%, the speed will be different in the two cases but in the same direction. If the power factor is less than 50%, the rotation will be in opposite directions in the two cases.

When instrument transformers are used, care must be exercised in determining correct connections; if terminals of similar instantaneous polarity have been marked on both current and voltage transformers, these connections can be verified and the usual test made to determine power factor. If the polarities have not been marked, or if the identities of instrument transformer leads have been lost in a conduit, the correct connections can still be established, but the procedure is more lengthy.

Use of Instrument Transformers with Watthour Meters. When the capacity of the circuit is over 200 A, instrument current transformers are generally used to step down the current to 5 A. If the voltage is over 480 V, current transformers are almost invariably employed, irrespective of the magnitude of the current, in order to insulate the meter from the line; in such cases, voltage transformers are also used to reduce the voltage to 120 V. Transformer polarity markings must be observed for correct registration. The ratio and phase-angle errors of these transformers must be taken into account where high accuracy is important, as in the case of a large installation. These errors can be largely compensated for by adjusting the meter speed.

Reactive Voltampere-Hour (Var-Hour) Meters. Reactive voltampere-hour (var-hour) meters are generally ordinary watthour meters in which the current coil is inserted in series with the load in the usual manner while the voltage coil is arranged to receive a voltage in quadrature with the load voltage. In 2-phase circuits, this is easily accomplished by using two meters as in power measurements, with the current coils connected directly in series with those of the “active” meters but with the voltage coils connected across the quadrature phases. Evidently, if the meters are connected to rotate forward for an inductive load, they will rotate backward for capacitive loads. For 3-phase 3-wire circuits and 3-phase 4-wire circuits, phase-shifting transformers are used normally and complex connections result.

Errors of Var-Hour Meters. The 2- and 3-phase arrangements described above give correct values of reactive energy when the voltages and currents are balanced. The 2-phase arrangement still gives correct values for unbalanced currents but will be in error if the voltages are unbalanced. Both 3-phase arrangements give erroneous readings for unbalanced currents or voltages; an autotransformer arrangement usually will show less error for a given condition of unbalance than the simple arrangement with interchanged potential coils.

Total var-hours, or “apparent energy” expended in a load, is of interest to engineers because it determines the heating of generating, transmitting, and distributing equipment and hence their rating and investment cost. The apparent energy may be computed if the power factor is constant, from the observed watthours P and the observed reactive var-hours Q; thus var-hours = $\sqrt{P^2 + Q^2}$. This method may be greatly in error when the power factor is not constant; the computed value is always too small.

A number of devices have been offered for the direct measurement of the apparent energy. In one class *(a)* are those in which the meter power factor is made more or less equal to the line power factor. This is accomplished automatically (in the Angus meter) by inserting a movable member in the voltage-coil pole structure which shifts the resulting flux as line power factor changes. In others, autotransformers are used with the voltage elements to give a power factor in the meter close to expected line power factor. By using three such pairs of autotransformers and three complete polyphase watthour-meter elements operating on a single register, with the record determined by the meter running at the highest speed, an accuracy of about 1% is achieved, with power factors ranging from unity down to 40%. In the other class *(b)*, vector addition of active and reactive energies is accomplished either by electromagnetic means or by electromechanical means, many of them very ingenious. But the result obtained with the use of modern watthour and var-hour meters are generally adequate for most purposes.

The *accuracy* of a watthour meter is the percentage of the total energy passed through a meter which is registered by the dials. The watthours indicated by the meter in a given time are noted, while the actual watts are simultaneously measured with standard instruments. Because of the time required to get an accurate reading from the register, it is customary to count revolutions of the rotating element instead of the register. The accuracy of the gear-train ratio between the rotating element and the first dial of the register can be determined by count. Since the energy represented by one revolution, or the *watthour constant,* has been assigned by the manufacturer and marked on the meter, the indicated watthours will be $K_h \times R$, where K_h is the watthour constant and R the number of revolutions.

Reference Standards. *Reference standards* for dc meter tests in the laboratory may be ammeters and voltmeters, in portable or laboratory-standard types, or potentiometers; in ac meter tests, use is made of indicating wattmeters and a time reference standard such as a stopwatch, clock, or tuning-fork or crystal-controlled oscillator together with an electronic digital counter. A more common reference is a standard watthour meter, which is started and stopped automatically by light pulsing through the anticreep holes of the meter under test.

The *portable standard* watthour meter (often called *rotating standard*) method of watthour-meter testing is most often used because only one observer is required and it is more accurate with fluctuating loads. Rotating standards are watthour meters similar to regular meters, except that they are made with extra care, are usually provided with more than one current and one voltage range, and are portable. A pointer, attached directly to the shaft, moves over a dial divided into 100 parts so that fractions of a revolution are easily read. Such a standard meter is used by connecting it to measure the same energy as is being measured by the meter to be tested; the comparison is made by the "switch" method, in which the register only (in dc standards) or the entire moving element (in ac standards) is started at the beginning of a revolution of the meter under test, by means of a suitable switch, and stopped at the end of a given number of revolutions. The accuracy is determined by direct comparison of the number of whole revolutions of the meter under test with the revolutions (whole and fractional) of the standard. Another method of measuring speed of rotation in the laboratory is to use a tiny mirror on the rotating member which reflects a beam of light into a photoelectric cell; the resulting impulses may be recorded on a chronograph or used to define the period of operation of a synchronous electric clock, etc.

Watthour meters used with instrument transformers are usually checked as secondary meters; that is, the meter is removed from the transformer secondary circuits (current transformers must first be short-circuited) and checked as a 5-A 120-V meter in the usual manner. The meter accuracy is adjusted so that when the known corrections for ratio and phase-angle errors of the current and potential transformers have been applied, the combined accuracy will be as close to 100% as possible, at all load currents and power factors. An overall check is seldom required both because of the difficulty and because of the decreased accuracy as compared with the secondary check.

General precautions to be observed in testing watthour meters are as follows: (1) The test period should always be sufficiently long and a sufficiently large number of independent readings should be taken to ensure the desired accuracy. (2) Capacity of the standards should be so chosen that readings will be taken at reasonably high percentages of their capacity in order to make observational or scale errors as small as possible. (3) Where indicating instruments are used on a fluctuating load, their average deflections should be estimated in such a manner as to include the time of duration of each deflection as well as the magnitude. (4) Instruments should be so connected that neither the standards nor the meter being tested is measuring the voltage-circuit loss of the other, that the same voltage is impressed on both, and that the same load current passes through both. (5) When the meter under test has not been previously in circuit, sufficient time should be allowed for the temperature of the voltage circuit to become constant. (6) Guard against the effect of stray fields by locating the standards and arranging the temporary test wiring in a judicious manner.

Meter Constants. The following definitions of various meter constants are taken from the Code for Electricity Metering, 6th ed., ANSI C12.

Register constant K_r is the factor by which the register reading must be multiplied in order to provide proper consideration of the register or gear ratio and of the instrument-transformer ratios to obtain the registration in the desired units.

Register ratio R_r is the number of revolutions of the first gear of the register, for one revolution of the first dial pointer.

Watthour constant K_h is the registration expressed in watthours corresponding to one revolution of the rotor. (When a meter is used with instrument transformers, the watthour constant is expressed in terms of primary watthours. For a secondary test of such a meter, the constant is the primary watthour constant, divided by the product of the nominal ratios of transformation.)

Test current of a watthour meter is the current marked on the nameplate by the manufacturer (identified as TA on meters manufactured since 1960) and is the current in amperes which is used as the basis for adjusting and determining the percentage registration of a watthour meter at heavy and light loads.

Percentage registration of a meter is the ratio of the actual registration of the meter to the true value of the quantity measured in a given time, expressed as a percentage. Percentage registration is also sometimes referred to as the *accuracy* or *percentage accuracy* of a meter. The value of one revolution having been established by the manufacturer in the design of the meter, meter watthours = $K_h \times R$, where K_h is the watthour constant and R is the number of revolutions of rotor in S seconds. The corresponding power in meter watts is $P_m = (3600 \times R \times K_h)/S$. Hence, multiplying by 100 to convert to terms of percentage registration (accuracy),

$$\text{Percentage registration} = \frac{Kh \times R \times 3600 \times 100}{PS}$$

where P is true watts. This is the basic formula for watthour meters in terms of true watt reference.

Average Percentage Registration (Accuracy) of Watthour Meters. The Code for Electricity Metering makes the following statement under the heading, "Methods of Determination":

> The percentage registration of a watthour meter is, in general, different at light load than at heavy load, and may have still other values at intermediate loads. The determination of the average percentage registration of a watthour meter is not a simple matter as it involves the characteristics of the meter and the loading. Various methods are used to determine one figure which represents the average percentage registration, the method being prescribed by commissions in many cases. Two methods of determining the average percentage registration (commonly called "average accuracy" or "final average accuracy") are in common use:
>
> *Method 1.* Average percentage registration is the weighted average of the percentage registration at light load (*LL*) and at heavy load (*HL*), giving the heavy-load registration a weight of 4. By this method:
>
> $$\text{Weighted average percentage registration} = \frac{LL + 4HL}{5}$$
>
> *Method 2.* Average percentage registration is the average of the percentage registration at light load (*LL*) and at heavy load (*HL*). By this method:
>
> $$\text{Average percentage registration} = \frac{LL + HL}{2}$$

In-Service Performance Tests. *In-service performance tests,* as specified in the Code for Electricity Metering, ANSI C12, shall be made in accordance with a *periodic test schedule,* except that self-contained single-phase meters, self-contained polyphase meters, and 3-wire network meters also may be tested under either of two other systems, provided that all meters are tested under the same system. These systems are the *variable interval plan* and the *statistical sampling plan*.

The chief characteristic of the *periodic-internal system* is that a fixed percentage of the meters in service shall be tested annually. In the test intervals specified below, the word *years* means calendar years. The periods stated are recommended test intervals. There may be situations in which individual meters, groups of meters, or types of meters should be tested more frequently. In addition, because of the complexity of installations using instrument transformers and the importance of large loads, more frequent inspection and test of such installations may be desirable. In general, periodic test schedules should be as follows:

1. Meters with surge-proof magnets and without demand registers or pulse initiators—16 years.
2. Meters without surge-proof magnets and without demand registers or pulse initiators—8 years.

The chief weaknesses of the preceding periodic test schedule are that it fails to recognize the differences in accuracy characteristics of various types of meters as a result of technical advance in meter design and construction, and fails to provide incentives for maintenance and modernization programs.

The *variable interval plan* provides for the division of meters into homogeneous groups and the establishment of a testing rate for each group based on the results of in-service performance tests made on meters longest in service without test. The maximum test rate recommended is 25% per year. The minimum test rate recommended for the testing of a sufficient number of meters to provide adequate data to determine the test rate for the succeeding year. The provisions of the variable interval plan recognize the difference between various meter types and encourage adequate meter maintenance and replacement programs. See Section 8.1.8.5 of ANSI C12 for details of operation of this plan.

The *statistical sampling program* included is purposely not limited to a specific method, since it is recognized that there are many acceptable ways of achieving good results. The general provisions of the statistical sampling program provide for the division of meters into homogeneous groups, the annual selection and testing of a random sample of meters of each group, and the evaluation of the test results. The program provides for accelerated testing, maintenance, or replacement if the analysis of the sample test data indicates that a group of meters does not meet the performance criteria. See Section 8.1.8.6 of ANSI C12 for details of the operation of this program.

Ampere-Hour Meters. Ampere-hour meters measure only electrical quantity, that is, coulombs or ampere-hours, and therefore, where they are used in the measurement of electrical energy, the potential is assumed to remain constant at a "declared" value, and the meter is calibrated or adjusted accordingly.

Ampere-hour or volt-hour meters for alternating current are not practical but ampere-squared-hour or volt-squared-hour meters are readily built in the form of the induction watthour meter. Ampere-hours or volt-hours are then obtainable by extracting the square root of the registered quantities.

Maximum-Demand Meters. Some methods of selling energy involve the maximum amount which is taken by the customer in any period of a prescribed length, that is, the maximum demand. Many types of meters for measuring this demand have been developed, but space permits only a brief description of a few. There are two general classes of demand meters in common use: (1) integrated-demand meters and (2) thermal, logarithmic, or lagged-demand meters. Both have the same function, which is to meter energy in such a way that the registered value is a measure of the load as it affects the heating (and therefore the load-carrying capacity) of the electrical equipment.

Integrated-Demand Meters. Integrated-demand meters consist of an integrating meter element (kWh or kvarh) driving a mechanism in which a timing device returns the demand actuator to zero at the end of each timing interval, leaving the maximum demand indicated on a passive pointer, display, or chart, which in turn is manually reset to zero at each reading period, generally 1 month. Such demand mechanisms operate on what is known as the block-interval principle. There are three types of block-interval demand registers: (1) the indicating type, in which the maximum demand obtained between each reading period is indicated on a scale or numeric display, (2) the cumulative type, in which the accumulated total of maximum demand during the preceding periods is indicated during the period after the device has been reset and before it is again reset, that is, the maximum demand for any one period is equal or proportional to the difference between the accumulated readings before and after reset, and (3) the multiple-pointer form, in which the demand is obtained by reading the position of the multiple pointers relative to their scale markings. The multiple pointers are resettable to zero.

Another form of demand meter, usually in a separate housing from its associated watthour meter, is the recording type, in which the demand is transferred as a permanent record onto a tape by printing, punching, or magnetic means or onto a circular or strip chart. A special form of tape recording for demand metering that has come into wide use in recent years is the pulse recorder, in which pulses from a pulse initiator in the watthour meter are recorded on magnetic tape or punched paper tape in a form usable for machine translation by digital-data-processing techniques. Advantages of this system are its great flexibility, freedom from the operating difficulties inherent in inked charts, and freedom from many of the personal errors of manual reading and interpretation of charts.

Thermal, Logarithmic, or Lagged-Demand Meters. These are devices in which the indication of the maximum demand is subject to a characteristic time lag by either mechanical or thermal means. The indication is often designed to follow the exponential heating curve of electrical equipment. Such a response, inherent in thermal meters, averages on a logarithmic and continuous basis, which means that more recent loads are heavily weighted but that, as time passes, their effect decreases. The time characteristics for the lagged meter are defined as the nominal time required for 90% of the final indication with a constant load suddenly applied.

Concordance of Demand Meters and Registers. The measurement of demand may be obtained with meters and registers having various operating principles and employing various means of recording or indicating the demand. On a constant load of sufficient duration, accurate demand meters and registers of both classifications will give the same value of maximum demand, within the limits of tolerance specified. On varying loads, the values given by accurate meters and registers of different classifications may differ because of the different underlying principles of the meters themselves. In commercial practice, the demand of an installation or a system is given with acceptable accuracy by the record or indication of any accurate demand meter or register of acceptable type.

3.1.12 Electrical Recording Instruments

Recording instruments are, in many instances, essentially high-torque indicating instruments arranged so that a permanent, continuous record of the indication is made on a chart. They are made for recording all electrical quantities that can be measured with indicating instruments—current, voltage, power, frequency, etc. In general, the same type of electrical mechanism is used—permanent-magnet moving-coil for direct current and moving-iron or dynamometer for alternating current. The indicator is an inking pen or stylus that makes a record on a chart moving under it at constant speed. This requires a higher torque to overcome friction, so the operating power required for a recording instrument is greater than for a simple indicating instrument. Overshoot is generally undesirable, and recording instruments are slightly overdamped, whereas indicating instruments are usually somewhat underdamped. Some recorders use strip charts; graduations along the length of the chart are usually of time intervals, and the graduations across the chart represent the instrument scale. Alternatively, the chart may be circular, with radial graduations for the instrument scale and time markers around the circumference. The chart paper should be well made and glazed to minimize dimensional changes from temperature and humidity. The ink should be in accordance with the maker's specification for the particular paper used so that it is accepted readily and does not run or blot the paper. Chart drives may be electrical or clockwork. In strip charts, perforations along the edges of the paper are engaged by a drive pinion; circular charts are rotated from a central hub.

Potentiometric self-balancing recorders are systems incorporating dc potentiometers, used either alone or with a transducer to measure various quantities.

Transducers include those for voltage, current, power, power factor, frequency, temperature, humidity, steam or water flow, gas velocity, neutron density, and many other applications.

Types of systems are classified according to the means of detecting and correcting electrical unbalance in the potentiometer circuit.

Accuracy on the order of $^1/_4$% may be expected from potentiometer recorders. To maintain this accuracy, the potentiometer is referenced against a standard cell or a reference voltage provided by a Zener diode. This may be performed by the operator pressing a button to give manual standardization whenever desired. A further refinement is to have automatic standardization, in which the operation is intiated by the chart-drive motor at specified intervals.

Range extension of potentiometric recorders upward is by means of shunt or series resistors. Extension below the basic range of the recorder requires preamplifiers.

Measurement of ac quantities requires the use of ac-to-dc transducers, for example, thermocouples, rectifiers, etc.

Alternating-current potentiometer recorders are simpler than the dc types because they require no standardization against a standard cell or Zener reference voltage and ac-to-dc conversion is not required, eliminating the requirement for a vibrator or saturable reactor. The amplifier and motor-control circuits can be the same as in the dc recorder. By far, the greatest application is with ac bridges, where the ac amplifier acts as an unbalance detector. Strain-gage bridges and bridges which employ platinum or nickel resistive elements for narrow-range temperature measurements frequently employ recorders of this type.

Proximity-type recorders use a high-frequency oscillator whose operation is started or stopped by the insertion of a metal vane into a pair of coils. If the vane is mounted on the pointer of an indicating instrument, the oscillator can sense movement between the pointer and a pair of coils fitted to the oscillator. Servo motion of the coils on displacement of the instrument pointer is accomplished by coupling the oscillator output to the input of a servo amplifier which drives the control motor. This gives a graphic record that follows but does not constrain movement of the instrument pointer. In this way, quantities which can operate an indicating instrument can be recorded without using a transducer.

Telemetering is the indicating or recording of a quantity at a distant point. Telemetering is employed in power measurements to show at a central point the power loads at a number of distant stations and often to indicate total power on a single meter, but practically any electrical quantity which is measured can be transmitted, together with a large number of nonelectrical quantities such as levels, positions, and pressures. Telemetering systems may be classified by type: current, voltage, frequency, position, and impulse.

1. In *current* systems, the movement of the primary measuring element calls for a current in the attached control member to balance the torque created by the quantity measured. This balancing current (usually dc) is sent over the transmitting circuit to be indicated and recorded. Totalizing is possible by the addition of such currents from several sources in a common indicator. The receiver may be as much as 50 mi from the transmitter.
2. In *voltage* systems, a voltage balance may be produced through a control-member voltmeter, or a voltage may be generated by thermocouples heated by the quantity to be measured, or produced as an *IR* drop as a result of a current torque balance, or generated by a generator driven at a speed proportional to the measured quantity. These voltages, however produced, are recorded at a distance by a potentiometer recorder. Here, also, the recorder may be 50 mi from the transmitter.
3. A *variable frequency* may be produced for telemetering by causing the primary element to move a capacitor plate in an rf oscillator or to change the speed of a small dc motor driving an alternator. High-frequency systems cannot be used for transmission over many miles.
4. In *position* systems, the movement of the primary element or of a pilot controlled by the primary element is duplicated at a distance. The pilot may be a bridge balancing resistance or reactance, a variable mutual inductance, or a selsyn motor where the position of a rotor relative to a 3-phase stator is reproduced at the receiver end. Satisfactory operation is usually limited to a few miles.
5. The *impulse* type of transmission of measured quantities is represented by the largest number of devices. The number of impulses transmitted in a given time may represent the magnitude of the quantity being measured, and these may be integrated by a notching device or by a clutch, or the duration of the pulse may be governed by the primary element and interpreted at the receiver. If the impulses are transmitted at high frequency, inductance and capacitance effects in the transmitting line limit the distance of satisfactory transmission; systems using dc impulses operate over 50 to 250 mi.

3.1.13 Resistance Measurements

The SI unit of resistance, the ohm, has been determined directly in terms of the mechanical units by *absolute-ohm* experiments performed at the National Bureau of Standards and at national laboratories in other countries. The reactance of an inductor or capacitor of special construction whose value can be computed from its dimensional properties is compared with a resistance at a known frequency. The value of this resistance can then be assigned in absolute (or SI) units, in terms of length and time—the dimensions of the inductor or capacitor and the time interval corresponding to the comparison frequency. These measurements are made with high precision, and it is believed that the assigned value of the National Reference Standards of resistance, maintained at the NIST, differs from its intended *absolute* value by not more than 1 part in 10^6.

The *National Reference Standard* of resistance is a group of five 1-Ω resistors of special construction, sealed in double-walled enclosures containing dry nitrogen and kept in a constant temperature bath of mineral oil at 25°C at the NIST. To ensure that their values are constant, they are intercompared at least weekly, compared with other standards of differing construction quarterly, and compared with similar groups in other major national laboratories frequently. Absolute experiments to determine their SI values are performed at rather longer intervals because of the complexity of such experiments—a new experiment of this type may require 5 years or more to complete. This reference group serves as the basis for all resistance measurements made in the country.

Resistance standards, used in precise measurements, are made with high-resistivity metal, in the form of wire or strip. Manganin—a copper-nickel-manganese alloy—is generally used in resistance standards because, when properly treated and protected from air and moisture, it has a number of desirable characteristics, including stable value, low temperature coefficient, low thermal emf at junctions with copper, and relatively high resistivity. A copper-nickel-chromium-aluminum alloy, Evanohm, has been used for high-resistance standards, since it has the same desirable characteristics as manganin and a much higher resistivity.

Standards with nominal values exceeding a megohm (a million ohms) are generally of films of metals such as Nichrome, a nickel-chromium alloy, deposited on a glass substrate. Four forms of standard are in general use. The Thomas-type 1-Ω standard is widely used as a primary standard. The Reichsanstalt form was developed in the German National Laboratory; and the NIST form. All three are designed to be used with their current-terminal lugs in mercury cups and are generally suspended in an oil bath to dissipate heat and to hold the temperature constant at a known value during measurements.

The fourth type, in widespread use for secondary references and as a primary standard at the 10,000-Ω level, consists of one or more coils of Evanohm wound on mica cards or cylindrical formers and terminated in binding posts for use on benchtops. The primary standard version of this type of resistor generally has the resistance elements hermetically sealed in an oil-filled container which also contains some type of resistive temperature sensor.

For highest precision, power dissipation must be kept below 0.1 W (calibrations at the NIST are generally performed at 0.01 W), although as much as 1 W can be dissipated in stirred oil with very small changes in value. The maker's recommendations should be followed regarding safe operating current levels. High- and low-resistance standards use different terminal arrangements. In all standards of 1 Ω lower value and standards up to 10,000 Ω intended for use at the part-per-million (ppm) level of accuracy or better, the current and voltage terminals are separated, whereas in other standards they may not be. The four-terminal construction is required to define the resistance to be measured. Connections to the current-carrying circuit range from a few microhms upward and, in a two-terminal construction, would make the resistance value uncertain to the extent that the connection resistance varies. With four-terminal construction, the resistance of the standard can be *exactly* defined as the voltage drop between the voltage terminals for unit current in and out at the current terminals.

Current standards are precision four-terminal resistors used to measure current by measuring the voltage drop between the voltage terminals with current introduced at the current terminals. These standards, designed for use with potentiometers for precision current measurement, correspond in structure to the shunts used with millivoltmeters for current measurement with indicating instruments. Current standards must be designed to dissipate the heat they develop at rated current, with only a small temperature rise. They may be oil- or air-cooled, the latter design having much greater surface, since heat transfer to still air is much less efficient than to oil. An air-cooled current standard of 20 $\mu\Omega$ resistance and 2000-A capacity, has an accuracy of 0.04%. Very low resistance oil-cooled standards are mounted in individual oil-filled containers provided with copper coils through which cooling water is circulated and with propellers to provide continuous oil motion.

Alternating-current resistors for current measurement require further design consideration. For example, if the resistor is to be used for current-transformer calibration, its ac resistance must be identical with its dc resistance within 1/100% or better, and the voltage difference between its voltage terminals must be in phase with the current through it within a few tenths of a minute. Thin strips or tubes of resistance material are used to limit eddy currents and minimize "skin" effect, the current circuit must be arranged to have small self-inductance, and the leads from the voltage taps to the potential terminals should be arranged so that, as nearly as possible, the mutual inductance between the voltage and current circuits opposes and cancels the effect of the self-inductance of the current circuit. Figure 3-15 shows three types of construction. In (*a*) a metal strip has been folded into a very narrow U; in (*b*) the current circuit consists of coaxial tubes soldered together at one end to terminal blocks at the other end; in (*c*) a straight tube is used as the current circuit, and the potential leads are snugly fitting coaxial tubes soldered to the resistor tube at the desired separation and terminating at the center.

Resistance coils consist of insulated resistance wire wound on a bobbin or winding form, hard-soldered at the ends to copper terminal wires. Metal tubes are widely used as winding form for dc resistors because they dissipate heat more readily than insulating bobbins, but if the resistor is to be used in ac measurements, a ceramic winding form is greatly to be preferred because it contributes less to the phase-defect angle of the resistor. The resistance wire ordinarily is folded into a narrow loop and wound bifilar onto the form to minimize inductance. This construction results in considerable associated capacitance of high-resistance coils, for which the wire is quite long, and an alternative construction is to wind the coil inductively on a thin mica or plastic card. The capacitive effect is greatly reduced, and the inductance is still quite small if the card is thin.

Resistors in which the wire forms the warp of a woven ribbon have lower time constants than either the simple bifilar- or card-wound types. Manganin is the resistance material most generally employed, but Evanohm and similar alloys are beginning to be extensively used for very high resistance coils. Enamel or silk is used to insulate the wire, and the finished coil is ordinarily coated with shellac or varnish to protect the wire from the atmosphere. Such coatings do not completely exclude moisture, and dimensional changes of insulation with humidity will result in small resistance changes, particularly in high resistances where fine wire is used.

Resistance boxes usually have two to four decades of resistance so that with reasonable precision they cover a considerable range of resistance, adjustable in small steps. For convenience of connection, terminals of the individual resistors are brought to copper blocks or studs, which are connected into the circuit by means of plugs or of dial switches using rotary laminated brushes; clean, well-fitted plugs probably have lower resistance than dial switches but are much less convenient to use.

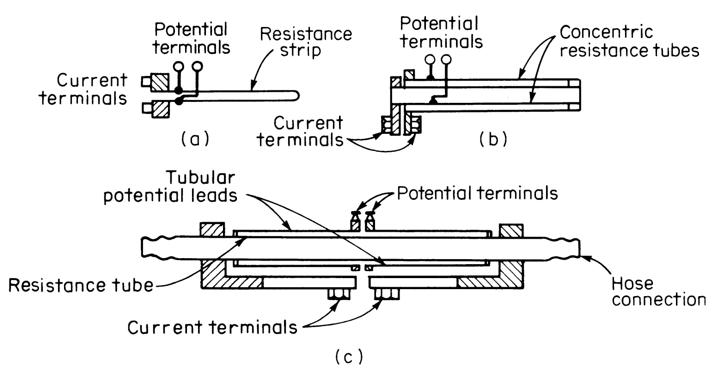


FIGURE 3-15 Types of low-inductance standard resistors.

The residual inductance of decade groups of coils due to switch wiring, and the capacitance of connected but inactive coils, will probably exceed the residuals of the coils themselves, and it is to be expected that the time constant of an assembly of coils in a decade box will be considerably greater than that of the individual coils.

Measurement of resistance is accomplished by a variety of methods, depending on the magnitude of the resistor and the accuracy required. Over the range from a few ohms to a megohm or more, an ohmmeter may be used for an accuracy of a few percent. A simple ohmmeter may consist of a milliammeter, dry cell, and resistor in a series circuit, the instrument scale being marked in resistance units. For a better value, the voltage drop is measured across the resistor for a measured or known current through it. Here, accuracy is limited by the instrument scales unless a potentiometer is used for the current and voltage measurements. The approach is also taken in the wide variety of digital multimeters now in common use. Their manufacturers' specifications indicate a range of accuracies from a few percent to 10 ppm (0.001%) or better from the simplest to the most precise meters. Bridge methods can have the highest accuracy, both because they are null methods in which two or more ratios can be brought to equality and because the measurements can be made by comparison with accurately known standards. For two-terminal resistors, a Wheatstone bridge can be used; for four-terminal measurements, a Kelvin bridge or a current comparator bridge can be used. Bridges for either two- or four-terminal measurements also may be based on resistive dividers. Because of their extremely high input impedance, digital voltmeters may be used with standard resistors in unbalanced bridge circuits of high accuracy.

Digital multimeters are frequently used to make low-power measurements of resistors in the range between a few ohms and a hundred megohms or so. Resolution of such instruments varies from 1% of full scale to a part per million of full scale. These meters generally use a constant-current source with a known current controlled by comparing the voltage drop on an internal "standard" resistor to the emf produced by a Zener diode. The current is set at such a level as to make the meter direct-reading in terms of the displayed voltage; that is, the number displayed by the meter reflects the voltage drop across the resistor, but the decimal point is moved and the scale descriptor is displayed as appropriate. Multimeters typically use three or more fixed currents and several voltage ranges to produce seven or more decade ranges with the full-scale reading from 1.4 to 3.9 times the range. For example, on the 1000-Ω range, full scale may be 3,999.999 Ω. Power dissipated in the measured resistor generally does not exceed 30 mW and reaches that level only in the lowest ranges where resistors are usually designed to handle many times that power. The most accurate multimeters have a resolution of 1 to 10 ppm of range on all ranges above the 10-Ω range. Their sensitivity, linearity, and short-term stability make it possible to compare nominally equal resistors by substitution with an uncertainty 2 to 3 times the least count of the meter. This permits their use in making very accurate measurements, up to 10 ppm, or resistors whose values are close to those of standards at hand. Many less expensive multimeters have only two leads or terminals to use to make measurements. In those cases, the leads from the meter to the resistor to be measured become part of the measured resistance. For low resistances, the lead resistance must be measured and subtracted out, or zeroed out.

The *Wheatstone bridge* is generally used for two-terminal resistors. In the low-resistance range where four-terminal construction is normal, the resistance of connections into the network may be a significant fraction of the total resistance to be measured, and the Wheatstone network is not applicable. Figure 3-16 shows the arrangement of a Wheatstone bridge, where A, B, and C are known resistances, and D is the resistance to be measured. One or more of the known arms is adjusted until the galvanometer G indicates a null; then $D = B(C/A)$. In case D is inductive, the battery switch S_1 should be closed before the galvanometer key S_2 to protect the galvanometer from the initial transient current. In a common form of bridge, B is a decade resistance, adjustable in small steps, while C and A (the ratio arms of the bridge) can be altered to select ratios in powers of 10 from $C/A = 10^{-3}$ to 10^3. If the value of the unknown resistor is not very different from that of a known resistor, accuracy may be improved by substituting the known and unknown in turn into arm D and noting the difference in

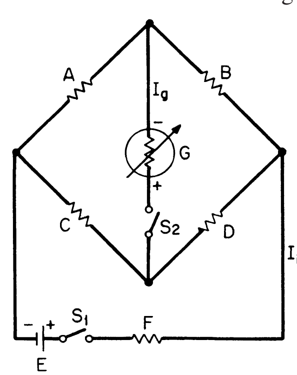


FIGURE 3-16 Wheatstone bridge.

balance readings of the adjustable arm *B*. Since there has been no change in the ratio arms, any errors they may have do not affect the difference measurement, and only those errors in arm *B* which were involved in the difference between the settings affect the difference value; in effect, the unknown is measured in terms of a known resistor by a substitution procedure. An alternative form of Wheatstone bridge is frequently assembled from standards and a ratio box of limited range called a *direct-reading ratio set*. This latter has a nominal ratio of unity, with ratio adjustments ranging from 1.005000 to 0.995000, that is, four decades of adjustment, of which the largest has steps of 0.1%. If a balance is made with the two standards in arms *B* and *D* and a second balance with the standards interchanged, their difference is half the difference between the balance readings. A similar technique can be used wherever small resistance differences are involved, for example, in the determination of temperature coefficients.

Bridge sensitivity can be determined in the following way. The voltage that would appear in the galvanometer branch of the circuit with switch S_2 open is

$$e = \frac{EBD\,\Delta B}{(B + D)^2}$$

where *E* is the supply voltage and ΔB is the amount in proportional parts by which *B* departs from balance. If, now, the voltage sensitivity of the galvanometer is known for operation in a circuit whose external resistance is that of the bridge as seen from the galvanometer terminals, its response for the unbalance ΔB can be computed. The current in the galvanometer with S_2 closed is

$$I_g = \frac{e}{G + BD/(B + D) + AC/(A + C)}$$

where *G* is the resistance of the galvanometer. If there is a large current-limiting resistance *F* in the battery branch of the bridge, the terminal voltage at the *AC* and *BD* junction points should be used rather than the supply voltage *E* in computing *e*. In connecting and operating a bridge, the allowable power dissipation of its components should first be checked to ensure that these limits are not exceeded, either in any element of the bridge itself or in the resistance to be measured.

Resistive voltage dividers can be used to form bridges for either two- or four-terminal resistance measurements. There are two common forms of resistive voltage divider—the Kelvin-Varley divider and the universal ratio set (URS)—with the former being the most commonly encountered. Each behaves as a potential divider with nearly constant input resistance and an open-circuit output potential of some rational fraction of the input, that fraction being given by the dial settings with calibration corrections applied. In the case of the Kelvin-Varley divider, the maximum ratio is 0.99999 . . . *X*, and outputs may be selected with a resolution as great as 1 part in 100 million of the input. Most Kelvin-Varley dividers have input resistances of 10,000 or 100,000 Ω. The URS was specifically designed to calibrate precision potentiometers. Its nominal input resistance is 2111.11 . . . 0 Ω, and that is also its full-scale dial designation. Its resolution, or one step of its least-significant dial, is either 1 or 0.1 mΩ. For bridge applications, either divider type appears as two adjacent (series-connected) bridge elements with a ratio of $r/(R - r)$, where *r* is the dial setting and *R* is the full-scale dial setting. In a Wheatstone or two-terminal type of bridge as shown in Fig. 3-16, the divider appears as resistors *A* and *B*, with *C* being the known resistor, or standard, and *D* being the unknown. In that case, the balance equation is

$$D/C = (R - r)/r$$

assuming that the low input of the divider is connected to the node between resistors *A* and *C* and its high input to the node between *B* and *D*.

Four-terminal applications are more complex, since four separate balances must be made to obtain the ratio between two resistors. The schematic is given in Fig. 3-17. To measure *B* in terms of *A*, the lead resistances between node pairs 1–2, 3–4, and 5–6, which we will call *x*, *y*, and *z*, respectively, must be eliminated. This is done by balancing the circuit with the resistor-side detector lead tied to each of the resistor potential leads at the terminals marked p_1, p_2, p_3, and p_4. The result is

$$\frac{A}{B} = \frac{r_2 - r_1}{r_4 - r_3}$$

where the *r*s are readings obtained by balancing the divider at each of the potential terminals.

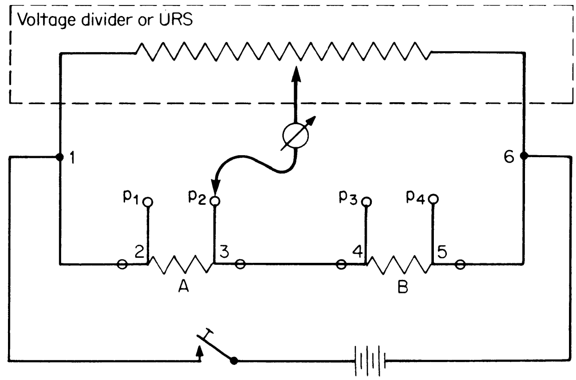


FIGURE 3-17 Four-terminal resistance measurements.

Both types of dividers must be calibrated. This can be done by comparison with a more accurate divider, dial by dial. Such a divider can readily be formed by using a number of nominally equal resistors in series. Each resistor is measured relative to the same standard and the results used to calculate the various ratios in the string of resistors. The string is then used to calibrate each setting of each dial in the voltage divider. In the case of the Kelvin-Varley divider, the dial corrections are interdependent; the correction for the steps in a particular dial depends on the settings of the less-significant dials.

Unbalanced bridge techniques have been made practical by the very high input resistances of modern digital instrumentation and are a satisfactory approach to resistance measurements when the values of the resistors being measured do not differ significantly from one another. They are particularly useful in cases where a process, not expected to change significantly, is being monitored using resistive sensors such as thermistors or copper or nickel resistors. The simplest case is that of a Wheatstone bridge such as that shown in Fig. 3-16. In it, the galvanometer *G* would be replaced by a digital meter of suitable sensitivity and sufficiently high input impedance to make bridge loading errors insignificant. The bridge relationship then becomes

$$\frac{B}{A + B} + \frac{v}{V} = C + D$$

where V is the voltage applied to the bridge, or $(E - I_b F)$, and v is the reading of the digital meter. In practice, the meter is generally used to measure V as well as v. If the individual elements of the resistor pairs $A - B$ and $C - D$ are nearly equal, the bridge is nearly at balance, v is small, and measurements of v and V need not be made at high accuracies. Resolution is not generally a problem for resistance element values of 100 Ω and higher because digital meters with least counts of 0.1 and 1 μV microvolts are commonly available.

A special form of Wheatstone bridge, known as a *Mueller bridge,* is commonly used for four-terminal measurements of the resistance of platinum resistance thermometers (PRTs). In this bridge, shown in Fig. 3-18, the effects of lead resistance of the PRT are eliminated by including two of the leads in adjacent bridge arms and making a second measurement after transposing the leads. The equations are

$$R_1 + l_1 = R_x + l_4 + S_1 \tag{a}$$

$$R_2 + l_4 = R_x + l_1 + S_2 \tag{b}$$

because the bridge is always used with the ratio arms *A* and *B* adjusted to be equal. These two equations may be added to eliminate lead resistances and result in the equation

$$R_x = 0.5\,(R_1 - S_1 + R_2 - S_2)$$

where R_1, S_1, R_2, and S_2 are the dial readings (with corrections applied) for conditions *A*, *B*.

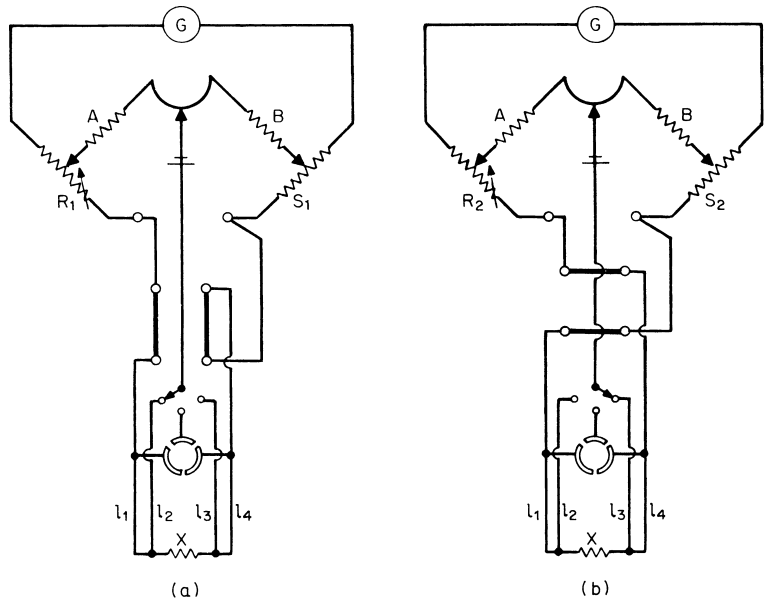


FIGURE 3-18 Mueller bridge.

There is increasing use of low-frequency square- and sinusoidal-wave bridges for PRT measurements. These bridges rely on the inherent ratio stability and accuracy of specially designed transformers and the increased sensitivity available with ac amplifiers to provide accuracies rivaling or surpassing those of the best dc bridges while requiring a minimum amount of upkeep. Both manual and automatic balancing types are available. Many contain one or more resistance reference standards kept at constant temperature in ovens. The transfer accuracies (i.e., accuracy available immediately after the bridge reference resistor has been calibrated) are very nearly equal to their least count, generally 0.1 ppm or better. Such bridges operate at 400 Hz or less to reduce problems with quadrature balances in the resistance being measured and its leads. They usually cover the range below 100 Ω.

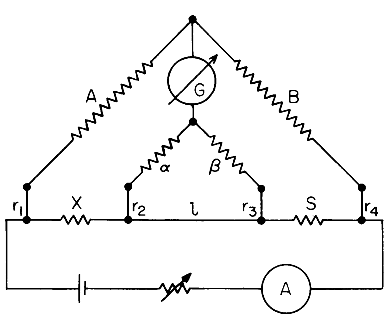


FIGURE 3-19 Kelvin double bridge.

The *Kelvin double bridge* is used for the measurement of low resistances of four-terminal construction, that is, whose current and voltage terminals are separate. Figure 3-19 shows the network. The balance equation is

$$\frac{X}{S} = \frac{A}{B} + \frac{1}{S}\frac{\beta}{\alpha + \beta + 1}\left(\frac{A}{B} - \frac{\alpha}{\beta}\right)$$

If the resistances X and S being compared are small so that the resistance of the link l connecting them is comparable, the term of the balance equation involving l could be significant, but if the ratio A/B is equal to the ratio α/β, the correction term vanishes. This equality can be demonstrated, after the bridge is balanced, by opening the link l; if the inner and outer ratios are equal, the bridge will remain balanced. It should be noted that the resistance of the leads r_1, r_2, r_3, and r_4 between the bridge terminals and the voltage terminals of the resistors may contribute to a ratio unbalance; these lead

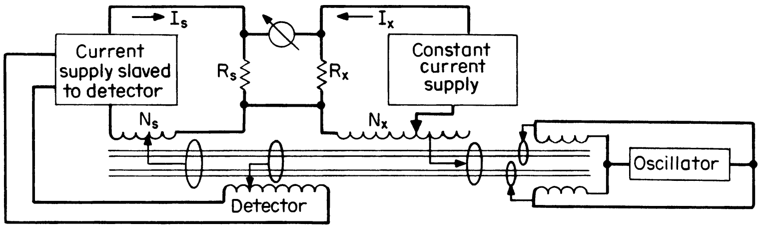


FIGURE 3-20 Current-comparator bridge.

resistances should be in the same ratio as the arms to which they are connected. In some Kelvin bridges, small adjustable resistors are provided for balancing leads; another technique is to shunt the α or β arm with a high resistance until $A/B = \alpha/\beta$ with the link removed. When this balance is achieved, the link l is replaced, and the main bridge balance is readjusted. In some bridges, the outer and inner ratio arms are adjustable only in decimal steps, and the main balance is secured by means of an adjustable standard resistor consisting of a Manganin strip with nine voltage taps of 0.01 or 0.001 Ω each and a Manganin slide wire. Portable bridges may use slide-wire arms and reference resistors to cover a range from 10 $\mu\Omega$ to 10 Ω.

In the *current-comparator bridge,* shown schematically in Fig. 3-20, the ratio of resistor currents is evaluated in the comparator as a balance of ampere-turns in the two circuits. $I_xN_x = I_sN_s$, so $R_x/R_s = N_x/N_s$ when $I_sR_s = I_xR_x$. A resistance determination that depends on the evaluation of a ratio is limited by the stability of that ratio. In Wheatstone and Kelvin bridges, the stability of individual resistors sets that limit; in the current-comparator bridge, the ratio is that of windings on a common magnetic core and therefore stable.

Since this bridge operates in terms of a ratio of currents for equal voltage drops, it can be used to determine power coefficients of low-value resistors. In a Kelvin bridge, the ratio of power dissipated is $P_s/P_x = R_s/R_x$ in the resistors compared; in the comparator bridge, this ratio is $P_s/P_x = R_x/R_s$. Thus, a low-value resistor operated at a substantial power level can be compared directly with a standard of higher resistance operated at a low power level.

Insulation resistance is generally measured by deflection methods. In the case of resistances on the order of a few megohms, a Wheatstone bridge may be used with low to moderate accuracy. A portable megohm bridge is made by General Radio Company. It operates as a Wheatstone bridge with an amplifier and dc indicating instrument as the detector system. A choice of high-resistance ratio arms gives ranges of 0.1 to 10^4 MΩ. On the highest range, the resolution limit is about 10^6 MΩ. The deflection methods fall in two general classes: (1) direct-deflection methods and (2) loss-of-charge methods.

Direct-deflection methods (insulation resistance) involve the simple application of Ohm's law. When the resistance is on the order of 1 MΩ, an ordinary voltmeter can give results that are good enough for most purposes. Two readings are taken, one with the voltmeter directly across the source of voltage, the other with the resistance to be measured connected in series with the voltmeter. The resistance is $R = V(d_1 - d_2)/d_2$, where V is the resistance of the voltmeter, d_1 is the voltmeter deflection on the first reading, and d_2 is the deflection on the second reading. The greater the resistance of the voltmeter per volt, the higher is the resistance that can be measured. For higher resistances, a reflecting galvanometer with high current sensitivity is used. Figure 3-21 is a diagram of the arrangement for measuring the insulation resistance of a cable.

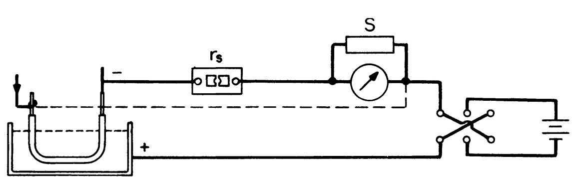


FIGURE 3-21 Diagram for measurement of insulation resistance of cable.

The measurement is made as follows: The galvanometer shunt S is set at the highest shunting value, and the circuit is closed. The shunt is decreased until a large, readable deflection is obtained. The reading is taken 1 min after closing the main switch. This procedure is repeated with only the standard resistor r_s (usually 0.1 or 1 MΩ) in the circuit, the specimen being short-circuited. The resistance of the specimen in megohms is $R = (G/d_1 s_1) - r_s$, where d_1 is the first reading and s_1 the multiplier corresponding to the shunt setting. G, the galvanometer megohm constant, is obtained from the second reading, $G = dr_s s$, where d is deflection, r_s the value of the standard resistor in megohms, and s the shunt multiplier. The conductor is preferably negative to the sheath or water. The standard resistor r_s is left in the circuit as a protection to the galvanometer against accidental short circuit in the sample. The guard for the cable ends is shown by the broken line. Removing braid for several inches at the ends of the sample and dipping the ends in hot paraffin tend to reduce leakage across the face of the insulation from sheath to central conductor, especially in damp weather.

The *loss-of-charge method* of measuring insulation resistance may be used when the resistance is very high, such as the resistance of porcelain and glass and the surface leakage resistance of line insulators. The principle is shown in Fig. 3-22, where the resistance r to be measured is connected in parallel with a capacitor C. Key a is closed and immediately opened, charging the capacitor. Key b is closed immediately after a is opened and the ballistic throw d_1 of the galvanometer noted. The process is repeated, but now a time t s is allowed to pass from the instant of charging before key b is closed and a deflection d_2 observed. The resistance is

$$r = \frac{t}{2.303C \log_{10}(d_1/d_2)} \qquad \text{M}\Omega$$

where C is the capacitance in microfarads.

The insulation resistance of the capacitor is not infinite and should be measured in a similar manner with r removed. The two resistances are in parallel, and the corrected value is

$$r = \frac{r_1 r_2}{r_2 - r_1}$$

where r_1 is the resistance value obtained in the first measurement and r_2 is the resistance of the capacitor. For even higher resistance, a growth-of-charge method may be used. In this case, the resistance to be measured is connected in series with a capacitor (preferably an air capacitor), and a known voltage E is applied for t s, the voltage on the capacitor being e at the end of this time. This value, e, is best measured with an electrostatic voltmeter connected continuously across C. The resistance is

$$r = \frac{t}{2.303C \log_{10}[E/(E - e)]} \qquad \text{M}\Omega$$

The *resistance of earth connections* may be measured by a three-electrode method. In Fig. 3-23, A is the connection whose resistance to earth is to be measured; it is temporarily disconnected from the distribution system while ground connection is preserved through a connection at D, either temporary or permanent. Two additional "grounds," B and C, are established, separated from each other and from A by not less than 15 ft. These auxiliary grounds may be pieces of metal buried in the earth,

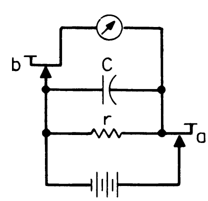


FIGURE 3-22 Leakage method of measuring insulation resistance.

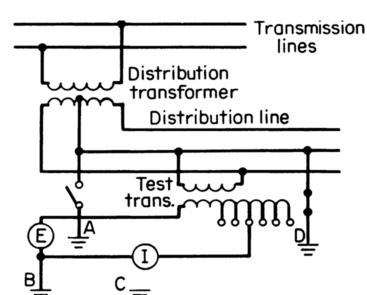


FIGURE 3-23 Resistance of earth connections.

such as a guy wire or a steel pole, making sufficient contact with ground for a good current reading. Resistances between the three electrodes taken in pairs are measured by a voltmeter-ammeter method. These resistances are r_{ab}, r_{bc}, and r_{ac}. Then the resistances are as follows:

$$\text{At } A: \quad R_a = \frac{r_{ab} - r_{bc} + r_{ac}}{2}$$

$$\text{At } B: \quad R_b = \frac{r_{ab} - r_{bc} - r_{ac}}{2}$$

$$\text{At } C: \quad R_c = \frac{r_{bc} - r_{ab} + r_{ac}}{2}$$

The measurement should be made with alternating current, which can be taken from the distribution system through an isolating transformer with secondary taps as shown. A low-range voltmeter is usually required. An Evershed ratio instrument is used for the measurement of ground resistance. One of the moving coils is traversed by the current sent through the ground from the attached hand-operated generator; the other is energized by the voltage drop to an auxiliary, driven electrode.

Faults in electric lines may be divided into two classes, closed- and open-circuit faults. *Closed-circuit faults* consist of *shorts,* where the insulation between conductors becomes faulty, and *grounds,* where the faulty insulation permits the conductor to make contact with the earth. *Open-circuit faults,* or *opens,* are produced by breaks in the conductors.

1. When the short is a low-resistance union of the two conductors, such as at *M* in Fig. 3-24, the resistance should be measured between the ends *AB*; from this value and the resistance per foot of conductor, the distance to the fault can be computed. A measurement of resistance between the other ends *A′B′* will confirm the first computation or will permit the elimination of the resistance in the fault, if this is not negligible.

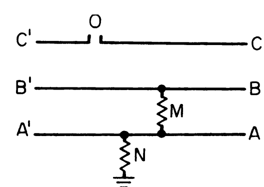


FIGURE 3-24 Line faults.

2. The location of a ground, as at *N* in Fig. 3-24, or of a high-resistance short is made by either of the two classic "loop" methods, provided that a good conductor remains.

Figure 3-25 shows the arrangement of the *Murray loop test*, which is suitable for low-resistance grounds. The faulty conductor and a good conductor are joined together at the far end, and a Wheatstone-bridge arrangement is set up at the near ends with two arms *a* and *b* comprising resistance boxes which can be varied at will; the two segments of line *x* and $y + l$ constitute the other two arms; the battery current flows through the ground; the galvanometer is across the near ends of the conductors. At balance,

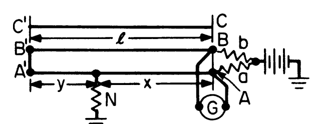


FIGURE 3-25 Murray loop.

$$\frac{a}{b} = \frac{x}{y + l} \quad \text{or} \quad \frac{a + b}{b} = \frac{x + y + l}{y + l} \quad \text{ohms}$$

The sum $x + y + l$ may be measured or known. If the conductors are uniform and alike and x and l are expressed as lengths, say, in feet,

$$x = \frac{2al}{a + b} \quad \text{ft}$$

If the ground is of high resistance, very little current will flow through the bridge with the arrangement of Fig. 3-25. In that case, battery and galvanometer should be interchanged, and thegalvanometer used should have a high resistance. If ratio arms a and b consist of a slide wire (preferably with extension coils), the sum $a + b$ is constant and the computation is facilitated. Observations should be taken with direct and reversed currents, especially in work with underground cables.

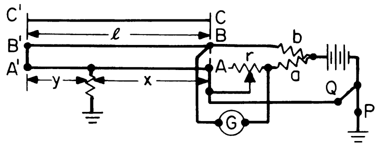


FIGURE 3-26 Varley loop.

In the *Varley loop,* shown in Fig. 3-26, fixed-ratio coils, equal in value, are employed, and the bridge is balanced by adding a resistance r to the near leg of the faulty conductor.

$$\frac{a}{b} = \frac{r + x}{y + l} \quad \text{or} \quad \frac{a + b}{b} = \frac{x + y + l + r}{y + l} \qquad \text{ohms}$$

If $a = b$,

$$x = y + l - r \quad \text{or} \quad x = \frac{1}{2(x + y + l - r)} \qquad \text{ohms}$$

The total line resistance $x + y + l$ is conveniently determined by shifting the battery connection from P to Q and making a new balance, r'. The equation then becomes $x = ½(r' - r)$. When a and b are slightly unequal, a second set of readings should be taken with a and b interchanged and the average values of r and r' substituted in the foregoing equations.

3. *Opens,* such as O in Fig. 3-24, are located by measuring the electrostatic capacitance to ground (or to a good conductor) of the faulty conductor and of an identical good conductor; the position of the fault is determined from the ratio of the capacitances.
4. *Shorts* and *grounds* may be detected by sending through the defective conductor an alternating current of audible frequency, say 1000 Hz. A pickup coil connected to a telephone receiver worn on the head of the tester is then carried along the line; the note in the receiver will cease when the fault has been passed.

3.1.14 Inductance Measurements

The *self-inductance,* or *coefficient of self-induction,* of a circuit is the constant by which the time rate of change of the current in the circuit must be multiplied to give the self-induced counter emf. Similarly, the *mutual inductance* between two circuits is the constant by which the time rate of change of current in either circuit must be multiplied to give the emf thereby induced in the other circuit. Self-inductance and mutual inductance depend upon the shape and dimensions of the circuits, the number of turns, and the nature of the surrounding medium.

Computable standards of self- or mutual inductance have been used traditionally in *absolute-ohm* determinations, but they are not suitable for use in assigning the values of other inductors—they are bulky and have relatively large capacitance to ground and considerable coupling to other circuits, their ratio of inductance to resistance is relatively low, and they exhibit appreciable skin effect even at moderately high frequencies, since they must be wound with rather heavy wire. All these undesirable features inevitably follow from the requirement that their values be computable from measured dimensions. Computable self-inductors and the primaries of computable mutual inductors are wound as single-layer solenoids on a dimensionally stable nonmagnetic form. The best of them are on cylinders of fused silica, and the winding is laid in a groove lapped into the form to ensure uniform winding pitch. The primary winding of a computable mutual inductor is in two or three sections spaced at such intervals that there is a region outside and in its central plane in which its field gradients are very small. The secondary—a multilayer winding—is located in this position so that its position and dimensions will be less critical.

Working standards of inductance are usually multilayer coils wound on nonmagnetic forms of Bakelite, marble, or ceramic to ensure reasonable dimensional stability. A toroidal core gives a coil that is practically immune to external magnetic fields. Approximate astaticism is also achieved by using two equal coils, connected in series and so located with respect to each other that their coupling with external fields tends to cancel each other. Since there is always capacitance associated with a winding, the effective value of an inductor will always be a function of frequency to a greater or lesser extent, and an accurate statement of value must necessarily include the frequency with which the value is associated. Inductance standards for radio frequencies are wound on open frames.

Single-layer winding or "loose basket weave" is essential to reduce the distributed capacitance and the consequent change of effective inductance with frequency. Insulating material is kept to a minimum to reduce dielectric loss.

Inductometers are continuously adjustable inductance standards. The Ayrton-Perry inductometer uses pairs of coaxial coils wound on zones of spheres; the outer pair is fixed, and the inner pair can be rotated about a vertical axis. The coils are so proportional that the scale is uniform over most of its length. This inductometer is not astatic, and its coupling with external fields can cause significant measurement errors. The Brooks inductometer, a better design from several viewpoints, consists of six link-shaped coils. The four stator coils are mounted in pairs above and below the rotor coils, which are located diametrically opposite one another in a flat disk. These two fixed- and moving-coil combinations are so connected that their coupling with external fields tends to cancel. The shape of the link coils gives a scale that is completely uniform except at its extreme ends, and the time constant of the inductometer is much higher than in the Ayrton-Perry arrangement. Ratio of maximum to minimum inductance is about 8:1, and change of calibration with wear in the bearings is negligibly small. Terminals of the fixed and movable coils are usually brought out separately so that inductometers can be used as either adjustable self-inductors or adjustable mutual inductors.

Measurement methods at power and audio frequency are (1) null methods employing bridges if accurate values are required or (2) deflection methods in which the inductance is computed from measured values of impedance and power factor, the measurements being made with indicating instruments—ammeter, voltmeter, wattmeter. At radio frequencies, resonance methods are used.

Bridges for inductance measurements can assume a variety of forms, depending on available components and reference standards, magnitude and time constant of the inductance to be measured, and a variety of other factors. In a four-arm bridge similar to the Wheatstone network, an inductance can be (1) compared with another inductance in an adjacent arm with two resistors forming the "ratio" arms or (2) measured in terms of a combination of resistance and capacitance in the opposite arm with two resistors as the "product" arms. It is generally better, where possible, to measure inductance in terms of capacitance and resistance rather than by comparison with another inductance because the problems of stray fields and coupling between bridge components are more easily avoided. The basic circuits will be described for a few bridges which can be used to measure inductance, but a more detailed reference should be consulted for a discussion of shielding, physical arrangement of components, effects of residuals, etc. In the balance equations which will be stated below, the inductance, L or M, will be expressed in henrys, the resistance R in ohms, capacitance in farads, and ω is $2\pi \times$ frequency in hertz. The time constant of an inductor is L/R; its storage factor Q is $\omega L/R$.

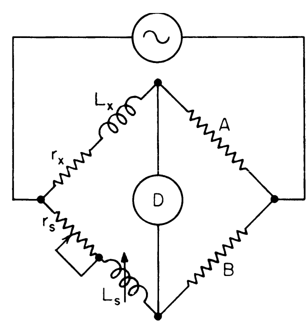


FIGURE 3-27 Inductance bridge.

Inductance comparison is accomplished in the simple Wheatstone network shown in Fig. 3-27, in which A and B are resistive ratio arms, L_x and r_x represent the inductor and the associated resistance being measured, and L_s and r_s are the reference inductor and the associated resistance (including that of the inductor itself) required to make the time constants of the two inductive arms equal. At balance,

$$\frac{A}{B} = \frac{L_x}{L_s} = \frac{r_x}{r_s}$$

An inductometer may be used to achieve balance, together with an adjustable resistance in the same bridge arm, as indicated in the diagram. If only a fixed-value standard inductor is available, balance can be secured by varying one of the ratio arms, but there also must be an adjustable resistance in series with L_x or L_s to balance the time constants of the inductive arms. Care must be taken to ensure that there is no inductive coupling between L_s and L_x, since this would lead to a measurement error.

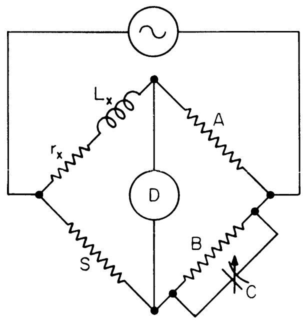


FIGURE 3-28 Maxwell-Wien inductance-capacitance bridge.

The *Maxwell-Wien bridge* for the determination of inductance in terms of capacitance and resistance is shown in Fig. 3-28. The balance equations are $L_x = ASC$ and $r_x = AS/B$. This bridge is widely used for accurate inductance measurements. It is most easily balanced by adjustments of capacitor C and resistor B; these elements are in quadrature, and therefore their adjustments do not interact.

Anderson's bridge, shown in Fig. 3-29, can be used for measurement over a wide range of inductances with reasonable values of R and C. Its balance equations are $L_x = CAS\,(1 + R/S + R/B)$ and $r_x = AS/B$. Balance adjustments are best made with R and r_x. This bridge also has been used to measure the residuals of resistors, a substitution method being employed in which the unknown and a loop of resistance wire with calculable residuals are substituted in turn into the L arm. If A and B are equal and if the resistances of the unknown and the calculable loop are matched, the residuals in the various bridge arms do not enter the final calculation, except the residual of Δr_x, the change in r_x between balances. The elimination of bridge-arm residuals from the *exact* balance equations is characteristic of substitution methods, and quite generally, residuals or corrections to the arms that are unchanged between the balances do not have to be taken into account in the final calculation when the difference is small between the substituted quantities.

Owen's bridge, shown in Fig. 3-30, can be used to measure a wide range of inductance with a standard capacitor C_b of fixed value, by varying the resistance arms S and A. In operation, the resistance S and capacitor $C_b(r_b)$ are usually fixed, balance being secured by successive adjustments of A and R. At balance,

$$r_x + R = (C_b/C_a)S + \omega L_x \omega C_b r_b$$

and $L_x(1 + \tan\delta_b \tan\delta_x) = C_b S(A + r_a)$. If $C_b(r_b)$ is a loss-free air capacitor so that $r_b = 0$ and $\tan\delta_b = 0$, $r_x = (C_b/C_z)S' - R$ and $L_x = C_b S(A + r_a)$. This is a bridge which is much used for examining the properties of magnetic materials; inductance may be measured with direct current superposed.With a low-reactance blocking capacitor in series with the detector and another in series with the source, a dc supply may be connected across the test inductance without current resulting in any other branch of the network; a high-reactance, low-resistance "choke" coil should be connected in series with the dc source.

Mutual inductance can be measured readily if an adjustable standard of proper range is available. Connections are made so the range of measurement is limited to values that can be read with the desired precision. Care should be taken in arranging the circuit to avoid coupling between the mutual inductors.

Iron-cored inductors vary in value with frequency and current, so measurements must be made at known current and frequency; bridge methods can, of course, be adapted to this measurement, care being exercised to ensure that the current capacities of the various bridge components are not exceeded. In such a case, the waveform of the voltage drop across the circuit branch containing the inductor may not be sinusoidal, whereas that across the other side of the bridge, containing linear resistances and reactances, may be undistorted. Generally, a tuned detector should be used.

Resonance methods can be used to measure inductance at radio frequencies. A suitable source is used to establish an rf field whose wavelength is λ m. The inductance L_x (microhenrys) to be measured is placed in this field and connected to a calibrated variable capacitor through a thermocouple

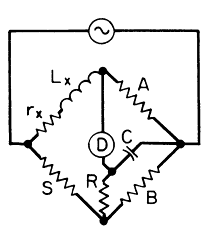


FIGURE 3-29 Anderson's bridge.

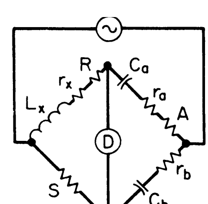


FIGURE 3-30 Owen's inductance-capacitance bridge.

ammeter (i.e., a current-indicating instrument without reactance). The capacitor is adjusted to resonance at a value of C (picofarads). Then, $L_x = 0.2815\lambda^2/C$. If a calibrated inductor L_s of the same order as L_x is available, the wavelength need not be known, and a substitution method can be used. The resonance settings are C_s and C_x, with L_s and L_x, respectively, in the circuit. Then, $L_x = L_s C_s / C_x$. The value of L_x is the effecitve inductance at the frequency of measurement and includes the effect of associated coil capacitance. The frequency of the source must not be affected by the substitution of L_x for L_s.

A *resonance-impedance method,* suitable for high frequencies, is indicated in Fig. 3-31. The capacitor C is adjusted untill the same current is indicated by the ammeter with switch K open or closed. (The applied voltage must be constant.) Then, $L_x = (1/2\omega^2 C)$ H if C is in farads and the frequency is $f = \omega/2\pi$. The waveform must be practically sinusoidal and the ammeter of negligible impedance. This method may be used to measure the effective inductance of choke coils with superposed direct current.

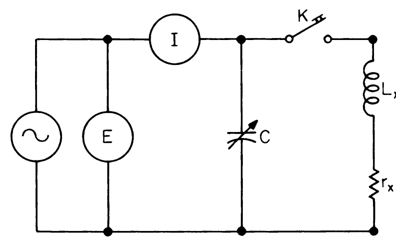


FIGURE 3-31 Reactance-impedance method of measuring inductance.

The *residual inductance* of a resistor or a length of cable at high frequency often can be determined by connecting the resistor in series with a fixed air capacitor and measuring its effective capacitance in an appropriate bridge with and without the series resistor S. If C_1 and C_2 are the measured capacitances in farads, without and with the series resistor, then

$$L = \frac{C_2 - C_1}{\omega^2 C_1 C_2} \qquad \text{and} \qquad S = \frac{1 - \omega^2 L C_1}{\omega C_1} \tan\delta$$

S is the effective resistance in ohms. L is the residual inductance in henrys, and δ is the loss angle of the capacitor-resistor combination computed from the second bridge balance.

3.1.15 Capacitance Measurements

The *capacitance* between two electrodes may be defined for measurement purposes as the charge stored per unit potential difference between them. It depends on their area, spacing, and the character of the dielectric material or materials, which is affected by the electric field between them. The value of a capacitor, measured in farads or a convenient submultiple of this unit, will be influenced quite generally by temperature, pressure, or any ambient condition that changes the dimensions or spacing of the electrodes or the characteristics of the dielectric. The *dielectric constant* of a material is defined as the ratio of the capacitance of a pair of electrodes, with the material occupying all the space affected by the field between them, to the capacitance of the same electrode configuration in vacuum.

Computable capacitors known to 1 part in 10^6 or better have been constructed at the NIST and at certain other national laboratories as a basis for their *absolute-ohm* determinations. Such capacitors now serve as the "base" unit in assigning values to standard capacitors. The electrode arrangement of these computable capacitors conforms to the geometry prescribed in the *Thompson-Lampard theorem:* If four cylindrical conductors of arbitrary sections are assembled with their generators parallel to form a completely enclosed cylinder in such a way that the internal cross capacitances per unit length are equal, then in vacuum these cross-capacitances are each

$$\frac{\ln 2}{4\pi^2 \mu_0 V^2}$$

In the mksa system of electrical units, where μ_0 has the assigned value $4\pi \times 10^{-7}$ and V is the speed of light in vacuum in meters per second, this capacitance is in farads per meter. The capacitance of such a cross capacitor is about 2 pF/m. A practical realization of such a capacitor consists of four equal closely spaced cylindrical rods with their axes parallel and at the corners of a square. Arranged as a three-terminal capacitor and with end effect eliminated, its value can be computed as accurately as its effective length can be measured.

The capacitance of vacuum capacitors with electrodes of simple geometry can be computed approximately in a few cases: (1) Flat, parallel plates with guard ring, $C = 0.08854A/t$ pF, where A is area of the guarded plate in square centimeters and t is spacing in centimeters between electrodes; if dimensions are in inch units, $C = 0.2249A/t$. (2) Coaxial cylinders with guard cylinders at both ends, $C = 0.24161L/\log(R_2/R_1)$ pF for centimeter units, or $C = 0.6137L/\log(R_2/R_1)$ pF for inch units, where L is the length of the guarded cylinder, R_1 is the radius of the inner cylinder, and R_2 the radius of the outer cylinder. (3) Concentric spheres, where R_1 is the radius of the inner sphere and R_2 is the radius of the outer sphere, $C = 1.1127R_1R_2/(R_2 - R_1)$ pF for centimeter units, or $C = 2.8262R_1R_2/(R_2 - R_1)$ pF for inch units. These formulas give only approximate values because they assume no contributing field beyond the edges of the bounding surfaces and take no account of possible eccentricity, lack of parallelism of surfaces, finite width of gap between guard and working electrode, etc., all of which would require small correction terms.

Standard capacitors at levels up to 10^3 pF are generally of a multiple-parallel-plate variety with dry gas (air or nitrogen) as dielectric. Low-temperature coefficient is secured by use of Invar—a low-expansion alloy—as the electrode material and a good degree of stability is achieved by careful, strain-free mounting of fully annealed components and by hermetically sealing the unit. A very high degree of stability has been achieved in a solid-dielectric construction at the 10-pF level in which a disk of fused silica is provided with fired-on silver electrodes. Direct capacitance is through the interior of the disk between its parallel faces, and a silver coating on the cylindrical face acts as guard electrode and confines the field. Very narrow gaps at the edges of the disk between the guard and active electrodes, together with continuation of the shielding in the mounting arrangement, eliminate the possibility of any portion of the measured capacitance being through an outside path between the parallel-plate electrodes. The assembly is hermetically sealed in dry nitrogen, in a shock-resistant, resilient mounting together with a resistance thermometer so that temperature corrections can be accurately applied. Standards of this type have shown variations less than 1 part in 10^7 over a year interval. From 10^3 pF to 1μF, standard capacitors generally have clear mica as dielectric. The electrodes may be metal foils laid out between the mica sheets, the assembly impregnated with paraffin, and the excess wax squeezed out under high pressure. In an alternative construction, the mica sheets are silvered, assembled under pressure, and the assembly hermetically sealed. Neither construction is as stable with time as the lower-value air-dielectric units, and the mica units are characterized by low but appreciable loss angles, whereas the loss angle of the air-dielectric standards is negligible in almost all applications. Continuously adjustable air capacitors have two stacks of interleaved parallel metal plates, one stack being mounted to rotate on an axis. The maximum capacitance occurs when the fixed and movable plates completely overlap; the minimum, a small value but not zero, occurs 180° from this position.

A *three-terminal construction* is required if the value of the capacitor is to be definite and independent of its proximity to other objects. In a nominally two-terminal arrangement, each of the electrodes has some capacitance to surrounding objects or to ground which may depend on spacing and which actually forms a second capacitance circuit in parallel with the capacitor of interest, as will be seen from Fig. 3-32*a* and *b*. It is only in case *c*, where there is an actual third electrode which completely encloses the other two, that the value can be made definite and completely independent of any object or field outside the assembly. A second advantage of the three-terminal construction is that the direct capacitance between the two enclosed electrodes can be made loss-free, since the solid insulation required to support them mechanically can be in the auxiliary capacitances between the enclosing shield electrode and the shielded electrodes.

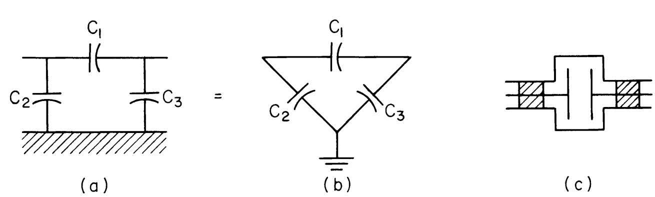


FIGURE 3-32 Two-terminal and three-terminal capacitors.

Methods of measuring capacitance can be classified as *null* methods, which quite generally involve the use of bridges, and *deflection* methods, in which some characteristic, usually impedance, is measured with the aid of indicating instruments. In the equations that follow, the capacitance C will be expressed in farads and resistance A, B, S in ohms. δ will be the *loss angle,* the amount which the current lacks of a true quadrature relation with voltage. The *power factor* of a capacitor is then cos $(\pi/2 - \delta) = \sin\delta$. The dissipation factor D is the name given to $\tan\delta$. It is convenient to represent a capacitor as consisting of a capacitance C (farads) in series with a resistance r (ohms) such that $\tan\delta = 2\pi fCr$ at a frequency f. The *power loss,* for an impressed voltage E (volts), is $P = 2\pi fCE^2 \sin\delta$. Since most bridges yield $\tan\delta$, the power loss can be expressed conveniently as $P = 2\pi fCE^2 \tan\delta$, where δ is small, or $P = \omega CE^2 D$.

Bridge methods for the comparison of capacitors are to be preferred over methods in which capacitance is determined in terms of inductance, since it is simple to shield capacitors so that their values are completely independent of neighboring objects and their electric fields are completely confined, whereas the magnetic fields of inductors cannot be so confined. Error voltages can enter bridges through coupling of an inductor with an external field, through mutual coupling with eddy-current circuits induced by the inductor in neighboring metal objects, etc.

DeSauty's bridge, shown in Fig. 3-33, is a simple Wheatstone network in which capacitors may be compared in terms of a resistance ratio. It should be noted that the loss angles of the two capacitance arms must be equal, so a series resistor is inserted in the branch with the smaller loss angle. In the case illustrated, the resistance S is in series with the reference capacitor C_s. At balance, $C_x = C_s(B/A)$, and $\tan\delta_x = \omega C_x r_x = \omega C_s(r_s + S) = \tan\delta_s + \omega C_s S$.

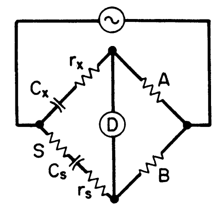


FIGURE 3-33 DeSauty's bridge.

Schering's bridge, shown in Fig. 3-34, has found wide application in measuring the loss angles of high-voltage power cables and high-voltage insulators. For this purpose, the supply voltage is connected as shown, and a ground connection is made at the junction of branches A and B so that the balance adjustments may be made close to ground potential. The adjustable components are generally A and C_p. It is also customary to enclose the A, B, and detector branches in a grounded screen and to protect this low-voltage section against possible breakdown of the test specimen by an air gap paralleling branch A. Such a gap can be set to spark over at 100 V or so, and provides a low-resistance path to ground for breakdown current from the specimen. The balance equations are $C_x = C_s(B/A)(1 + \tan\delta_s \tan\delta_p)$ and $\tan\delta_x = \omega C_p B + \tan\delta_s$. Usually, the reference capacitor C_s is a high-voltage air or compressed-gas capacitor with a negligible phase-defect angle, in which case the correction terms to the balancing equation drop out. The Schering bridge is also an excellent one to use for the comparison of capacitors at low voltage. For this purpose, it is used in its conjugate form with supply and detector branches interchanged to increase sensitivity. C_p must, of course, be connected across branch A instead of B if the loss angle of C_s is greater than that of C_x, with a corresponding modification of the balance equations. When the loss angles of C_s and C_x are both very small, adjustable capacitors must be connected across both A and B arms, and the difference in the phase-defect angles they introduce into the bridge must equal the difference in loss angles of C_s and C_x. This modification of the bridge is made necessary by the fact that the capacitance of an adjustable capacitor cannot be reduced to zero in the usual construction.

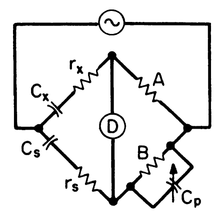


FIGURE 3-34 Schering's bridge.

The *transformer bridge* has been developed into the most precise tool available for the comparison of capacitors, especially for three-terminal capacitors with complete shielding. A three-winding transformer is used so that the bridge ratio is the ratio of the two secondary windings of the transformer which are of low resistance and uniformly distributed around a toroidal core to minimize leakage reactance. A stable ratio, known to better than 1 part in 10^7, can be achieved in this way.

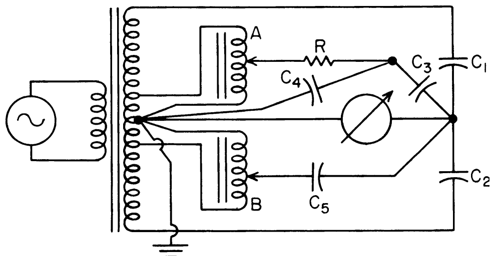


FIGURE 3-35 Transformer bridge for capacitor comparison.

A number of schemes for balancing adjustment have been used successfully. One of these, employing *inductive voltage dividers,* is shown schematically in Fig. 3-35 but simplified by omitting the necessary shielding. Current in phase with the main current is injected at the junction between the capacitors being compared, C_1 and C_2, to balance their inequality in magnitude. This current, through capacitor C_5, is controlled by adjusting the tap position on inductive voltage divider *B*, supplied from an appropriate tap point on the main transformer-ratio arm. Quadrature current, to balance the phase difference between C_1 and C_2, is similarly injected through *R* and the current divider $C_3/(C_3 + C_4)$, controlled by adjusting the tap point on divider *A*. The current divider is used so that *R* may have a reasonable value, a few megohms at most. In the illustrated network, it is assumed that $C_1 > C_2$ and that $\delta_1 < \delta_2$. The balance equations are $C_2 = C_1 + N_B C_5$ and $\delta_2 = \delta_1 + \omega R C_1 \cdot N_A \cdot C_3/(C_3 + C_4)$, where N_B is the fraction of the voltage across C_2 which is impressed on C_5, that is, the product of the tap-point ratios of the main transformer and divider B_5 and N_A is the corresponding fraction of the voltage across C_1 which is impressed on *R*. The reactance of C_3 and C_4 in parallel must be small compared with the resistance of *R*.

New automated impedance measurement instruments have come into being because of the ready availability of microprocessors. Some of these make use of the transformer techniques mentioned earlier, using relays to balance them by selecting ratios computed by the microprocessor from detector output voltages. Many have purely analog quadrature balance features. At least, one measures by passing the same current through the admittance to be measured and a reference resistor and computing the vector impedance of the unknown from the vector ratio of the voltage drops across it and the reference resistor. This is done using a 90° phase reference generated internally using digital synthesis techniques.

Many automated bridges are intended for testing of precision components over a broad range of frequencies and with programmable direct current or voltage biases. Their accuracies range from a few percent at high frequencies to 0.01% or better at audio frequencies. Their calibration is generally done using fixed-value two- or three-terminal or four-pair-terminal standards.

Detectors used in bridge measurements are selected with regard to frequency and impedance. *Vibration galvanometers* can be used at power frequencies in low-impedance circuits; they discriminate well against harmonics and have high sensitivity, but they must be tuned sharply to the use frequency.

Wave analyzers, which are commercially available with internal crystal control, also have a narrow passband and a high rejection of frequencies on either side. They can be used with a preamplifier when maximum sensitivity is required, and it is desirable that the preamplifier itself be sharply tuned in its first stage to improve noise rejection. This system can be used at any frequency throughout the audio region.

Cathode-ray oscilloscopes of adequate sensitivity (or used with tuned preamplifiers) make particularly good null detectors. If a phase-adjustable voltage from the bridge supply is impressed on the horizontal plates and the unbalance signal in the detector branch impressed on the vertical plates, the resulting Lissajous figure is an ellipse which, with proper phase adjustment, will change its opening with magnitude adjustment and the slope of its major axis with quadrature adjustment in the bridge. Balance is indicated by a straight horizontal trace on the screen. It is essential in this system that the initial stages of amplification be sharply tuned or that the bridge input be sinusoidal, for otherwise the pattern on the screen is confused and difficult to interpret. Phase discrimination of this type in the null detector is of considerable value in achieving balance, since it informs the operator of the individual magnitudes of inphase and quadrature unbalance.

Telephone receivers may be used at audio frequencies (maximum sensitivity being at about 1 kHz), but their response is usually quite broad, and the balance point may be masked by the presence of harmonics.

In *resonance methods* at radio frequencies, a thermocouple ammeter can be employed to show the current maximum. A crystal rectifier with an electronic voltmeter is used at ultrahigh frequencies.

Precautions in Bridge Measurements. The effect of stray magnetic fields can be minimized by using twisted-pair or coaxial leads and by avoiding loops in which an emf could be induced. Inductive coupling between bridge components should be avoided. Capacitive coupling existing between parts of the bridge which are not at the same potential will impress shunt capacitance across one or more of the bridge arms and modify the balance condition. Shielding must be used to minimize these effects.

Resonance methods are used for capacitance measurements at radio frequencies, a coil of known inductance L (microhenrys) at a known wavelength λ (m) being employed. Resonance is produced by varying λ and is detected with a thermocouple ammeter. At resonance, $C_x = 0.2815\lambda^2/L$ in picofarads. λ and L need not be known if a substitution method is used in which an adjustable capacitor with a range that includes C_x is connected in place of the unknown capacitor and adjusted to resonance without altering the frequency so that $C_x = C_s$. The leads used to connect the capacitors into the circuit must not be changed in length or position in making the substitution.

A *cavity resonator* can be used at frequencies on the order of 200 to 1000 MHz for measuring the characteristic of insulating material placed between electrodes within the cavity. Resonance is established with excitation of a small loop of wire within the cavity by connection to an oscillator, and resonance is shown by a crystal-rectifier probe connected to an electronic voltmeter.

3.1.16 Inductive Dividers

Inductive dividers are employed in precise voltage- and current-ratio applications. The ratios are used for comparing impedances and for calibrating devices with known nominal ratios such as other dividers, synchros, and resolvers. A divider usually consists of an autotransformer adjustable in decade steps. Such transformers, with high ratio accuracy for voltage or current comparison, have been made by using high-permeability magnetic-core materials and ingenious winding and connection techniques. Such a transformer can be represented electrically by the equivalent circuit of Fig. 3-36. The components of this circuit can be measured directly and will predict the performance of the divider. D is a perfect divider with infinite input impedance and zero output impedance. A' is the transfer ratio of D, the ratio of the voltage between the open-circuited *tap point* and the *low* end to the voltage between the *high* and *low* ends. A' is also the ratio of the short-circuit current between the high and low ends to the current into the tap point and out of the low point. Z_{oc} is impedance between high and low points, with the tap point open-circuited. This impedance is quite high and is a function of input voltage and frequency primarily. Its major components are the winding capacitance, charging inductance, and leakage reactance in parallel. Z_{sc} is the impedance between tap and low points, with the high and low points short-circuited together. This impedance is quite low and is a function of frequency and setting. Its major components are winding and contact resistances and the leakage inductance in series. The autotransformer configuration can produce voltage and current ratios of very high accuracy.

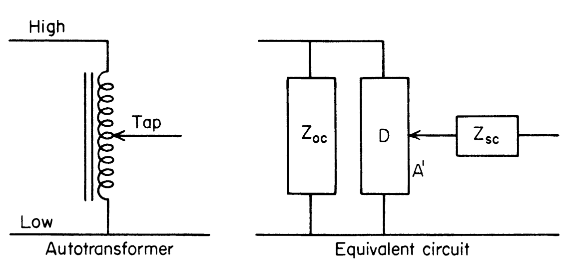


FIGURE 3-36 Autotransformer and equivalent circuit.

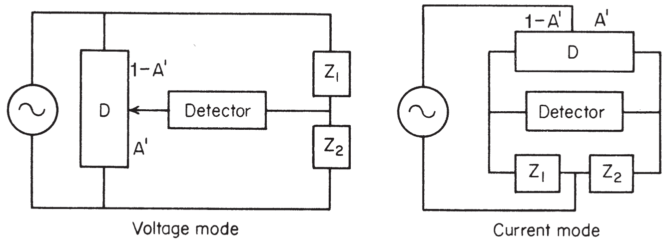


FIGURE 3-37 Voltage mode (*left*) and current mode (*right*) for impedance comparison

As a *voltage divider,* the circuit can be represented by a Thévenin equivalent consisting of a zero-impedance generator, with a voltage which is the product of the input voltage times the transfer ratio, and an output impedance equal to Z_{sc}. This low-output impedance provides high accuracy even with appreciable load admittance. For example, a 5000-Ω load will change the output voltage by only 0.1% if the output impedance is 5 Ω.

As a *current divider,* the circuit can be represented by Norton equivalent consisting of an infinite-impedence generator, with a current which is the product of the transfer ratio A' and the input current and an impedance equal to Z_{oc} in shunt across the output. This high-shunt output impedance provides high accuracy even with appreciable load impedance. For example, a 500-Ω load will receive a current only 0.1% less than short-circuit current if the output impedance Z_{oc} is 500,000 Ω.

Impedance comparison, using the comparison ratio $A'/(1 - A')$, can be accomplished in either the voltage mode of operation or the current mode of operation, as shown in Fig. 3-37. For impedance comparison, the divider impedances Z_{oc} and Z_{sc} are of no consequence. In the voltage-ratio mode, Z_{oc} is outside the bridge circuit, and at null no current is drawn through Z_{sc}. In the current-ratio mode, Z_{sc} is outside the bridge circuit, and at null there is no voltage across Z_{oc}. In either mode, the balance equation is $Z_2 = Z_1 A'/(1 - A')$.

3.1.17 Waveform Measurements

The instantaneous variations of current and voltage in a circuit can be measured by oscillographs, whose basic operating principle may be either that of a D'Arsonval galvanometer whose inertia is low enough to permit it to follow the variations or that of an electron beam which has no sensible inertia and whose deflection is governed by electric or magnetic fields. In addition to tracing waveforms, oscillographs are used for measurements of transient phenomena, such as those which occur in switching operations or in the impulse-voltage testing of insulating structures and disturbances resulting from lightning discharges. Transient phenomena also may be captured using digitizing oscilloscopes and transient digitizers (waveform recorders).

The *galvanometer oscillograph* may have a light low-inertia coil or, for higher-frequency response, a pair of thin metal ribbons tightly stretched across insulating bridges and tied together by a small mirror at their midpoints, mounted in the field of a permanent magnet. A light beam from the galvanometer mirror traces its response to varying current on a moving photographic film or, by means of an intermediate rotating mirror, on a stationary viewing screen. Galvanometer elements have been built with natural response frequencies as high as 8 kHz (a more common construction has a resonance frequency of about 3 kHz) and, if damped at about 0.7 of critical, have a response to signals which is practically free from distortion up to about half their resonant frequency; at resonant frequency, the deflection sensitivity has decreased to about 70% of their dc sensitivity for this damping.

3.1.18 Frequency Measurements

Reed-type frequency meters have a number of steel strips rigidly fastened to a bar at one end and free to vibrate at the other. These strips are located in the field of an electromagnet which is energized from the circuit whose frequency is to be measured. The strips have been accurately adjusted by solder weights to resonant vibration frequencies that differ by ¼ or ½ Hz, and the one with a period corresponding to the alternations of the voltage will be set into vibration.The free ends of the strips or reeds are turned up and painted white so that the reed which is vibrating will be indicated by an extended white band or blur.

The *Weston frequency meter* has fixed coils, 90° apart, and *a* movable element consisting of a simple, soft-iron core mounted on a shaft, with no control of any kind.

Resonant circuit meters, operating from circuits containing inductance and capacitance, can be made sensitive enough to indicate frequency variations of 0.01 Hz or less.

Precise frequency control is also accomplished with resonance techniques. Small-range indicators or recorders can be built as relays to monitor the frequency of a power system or generator, injecting an appropriate signal into a control system to restore frequency to a particular value. Such control may be made precise enough for use of the system frequency for electric-shock operation. Any tendency to frequency drift may be detected and corrected at the source by comparing an electric clock with a precise pendulum clock or one driven by a quartz-crystal oscillator.

Radio frequencies may be measured directly or indirectly. Direct measurement may be made with a wavemeter, an instrument with a tunable circuit and an ammeter to indicate the resonance frequency by a current maximum. In the indirect method, the unknown-frequency signal is introduced into a circuit with a precisely known frequency, and the beat frequency is counted. Quartz crystals maintained in temperature-regulated ovens will control the frequency of an oscillator to much better than 1 part in 10^6. Such a crystal-controlled oscillator, serving as a local reference standard of frequency, can be monitored against the very precise *standard* frequencies continuously broadcast by the NIST from its low-frequency station WWVL, operated at 60 kHz, or its high-frequency stations WWV and WWVH, which broadcast at a large number of higher frequencies. These broadcast frequencies are controlled by crystals operating under conditions that are most favorable to stability and are, in turn, monitored against the frequency of an atomic-beam resonator. The transmitted frequencies, as sent from the bureau stations, are accurate to about 1 part in 10^{12}. Frequencies from these broadcasts are modified somewhat in transmission by diurnal and moment-to-moment variations in the ionosphere, and their accuracy as received may be reduced by more than an order of magnitude.

Audio frequencies can be measured with a frequency-sensitive bridge, such as the Wien bridge with parallel- and series-connected capacitance-resistance arms, or they can be conveniently observed with a cathode-ray oscilloscope, if a known reference frequency is available. One set of plates of the oscilloscope is excited by the known- and the other set by the unknown-frequency signals. If the two frequencies have an exact, simple fractional relation, the Lissajous figure formed on the screen is stationary. For a 1/1 relationship, the pattern is an ellipse; for other fractional relationships, the pattern is more complicated, the relationship being determined from the number of loops. If the relationship cannot be represented by a simple fraction, the pattern will change continuously, and a count of the beat frequency is made over a measured time interval.

Electronic counters are widely used for frequency measurements. They work by counting the number of cycles of an input signal, or events, which occur in a very accurately known time interval (gate time). The gate time is based on the output of an internal standard oscillator (clock) or, optionally, on a reference-frequency signal input to the counter. Most counters of laboratory quality also can be used to measure the period of low-frequency signals, time intervals, the ratio of the frequencies of two input signals, and a total number of events. They also afford control of triggering, thus enabling the user to set trigger levels and slopes, noise rejection levels, and input attenuation levels. Output is via digital display, ranging from six to nine digits, and (usually) high-speed digital computer interface. Accuracies of frequency measurements are usually stated by the manufacturer to be ± clock accuracy ± 1 count. Most laboratory-grade counters can be equipped with high-stability crystal-based clocks, mounted in temperature-controlled ovens, and are stated to have drift rates as low as 2×10^{-8} per month. The frequency ranges covered are from nearly dc, directly or via period measurements, to as high as 500 MHz directly and to over 30 GHz using heterodyning techniques.

3.1.19 Slip Measurements

The *slip* of a rotating ac machine is the difference between its speed and the synchronous speed, divided by the synchronous speed; slip is usually expressed as a percentage. It may be computed from the measured speed of the machine and the synchronous speed, but direct methods are more accurate.

Millivoltmeter Method. If sufficient stray field is produced by the current in the secondary of an induction motor, a dc millivoltmeter connected to an adjacent coil of wire or across the motor shaft or frame will oscillate at slip frequency, each swing being one pole slip. In motors with wire-wound rotors, the millivoltmeter may be connected across the rotor slip rings.

Stroboscopic Method. In Fig. 3-38, a black disk with white sectors, equal in number to the number of poles of the induction motor, is attached to the induction-motor shaft. It is observed through another disk having an equal number of sector-shaped slits and carried on the shaft of a small self-starting synchronous motor, in turn fitted with a revolution counter which can be thrown in and out of gear at will. If n is the number of passages of the sectors, then $100n/n_s n_r$ = slip in percent, where n_s is the number of sectors and n_r is the number of revolutions recorded by the counter during the interval of observation. For large values of slip, the observations can be simplified by using only one sector ($n_s = 1$); then n = slip in revolutions.

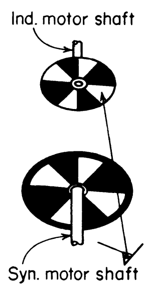


FIGURE 3-38 Slip measurement by stroboscopic method.

With a *synchronous light source* to illuminate the target on the induction-motor shaft, the synchronous motor is no longer necessary. An arc lamp connected across the ac supply may be used, but the carbons must be readjusted from time to time. A neon lamp makes a satisfactory source of light when the general illumination is not too bright. A portable stroboscope may consist of a gaseous discharge tube backed by a parabolic reflector to concentrate the light beam and an adjustable-frequency voltage source to trigger the flashlamp synchronously. The flash also can be triggered externally. Light output measured 1 m from the lamp may exceed 10^6 candela and flash duration may be as low as 0.5 μs.

Synchronizing. In order to connect any synchronous machine in parallel with another machine or system, the two voltages must be made equal and the machines must be synchronized, that is, the speed so adjusted that corresponding instantaneous values on the two waves are reached at the same instant, when they will be in exact phase. Furthermore, with polyphase machines, the direction of phase rotation must assuredly be the same. This, however, is usually made right once and for all when the machines are installed, the phases being so connected to the switches that the phase rotation will always be correct.

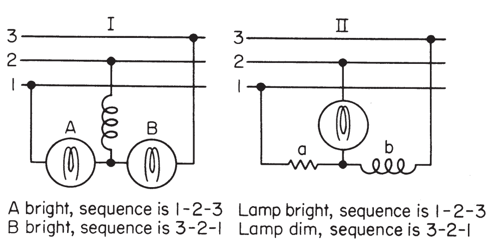


FIGURE 3-39 Phase-sequence indicators.

The *phase sequence of a 3-phase system* is often desired. Figure 3-39 shows two lamp methods. In I, two lamps and a highly reactive coil, such as the potential coil of a watthour meter, are used. The bright lamp indicates the particular phase sequence. In arrangement II, a noninductive resistance and a reactive coil of equal impedance are used in conjunction with a lamp, the brightness of which indicates the sequence.

3.1.20 Magnetic Measurements

The two classes of magnetic measurements are *field measurements,* such as the earth's field or the field in the air gap of a magnet, and *measurements to determine the characteristics* of magnetic materials.

Magnetic field measurements are commonly made by induction methods in which a coil is placed with its plane perpendicular to the field. Removing the coil to a point of zero flux or reversing the coil will induce in it an emf that can be measured by the ballistic deflection of a galvanometer in terms of its flux-linkage sensitivity (when operating in a circuit having the resistance of the search-coil circuit). In this measurement, $\int e\,dt = \mathrm{N}\,\Delta\phi 10^8$, where ϕ is the flux in maxwells enclosed by the coil and N is search-coil turns. The flux density B, in gauss, is ϕ/a, where a is the coil area in square centimeters.

The flux-linkage sensitivity of the galvanometer under the operating condition can be determined with the aid of a mutual inductor, with the galvanometer in the secondary circuit and a known current reversed in the inductor primary. Here, $\int e\,dt = 2MI$ volt-seconds, where M is mutual inductance in henrys and I is primary current in amperes. A Grassot-type fluxmeter can be used in place of a ballistic galvanometer. This is essentially a ballistic instrument in which restoring torque is reduced substantially to zero so that the deflection remains steady after the change in flux linkages.

Low field measurements are also made with magnetometers. This instrument uses a strip of high-permeability, low-coercive-force material (usually supermalloy) with an ac excitation coil that drives the material into saturation each half cycle at a frequency of a few kilohertz. A second-harmonic detector coil on the same strip will sense a bias field to which the assembly is exposed. A third coil on the strip supplies a measured offset ampere-turns to return the detector to zero, providing a very sensitive field measurement device. This is widely used in earth's field and other low-level field measurements. A portable flux-gate magnetometer, in which the vector-magnetic-field component at the sensor is neutralized by a current in a solenoid surrounding the sensor, has a resolution of 1 gamma at the neutralizing control. The magnitude (in gamma) of the neutralizing field is indicated on manually operated digital dials, and any difference between ambient field-vector component and neutralizing field is indicated on a meter whose range may be selected between 25 and 10^4 gammas. A nondirectional magnetometer system is based on proton gyromagnetic ratio and the functional relation between ambient field and resonance frequency in the sensor. This type of magnetometer is also used to sense small variations in the local earth's field.

Measurement of higher fields (20 to 20,000 G) and fields in spaces too confined for search coils are frequently made with Hall-effect gaussmeters. In a thin strip or film of a metal having a large Hall-effect coefficient and carrying a current, two points on opposite sides of the strip can be found between which there is no potential difference. If a magnetic field is then applied at a right angle to the plane of the strip, a potential will exist between these points which is proportional to the field. Germanium, bismuth, indium antimonide, and indium arsenide are the common materials for such probes, and they may be as small as 0.15×1.2 mm. Response of many of these instruments is fast enough to allow operation up to midrange audio frequencies.

In another type of gaussmeter, a small permanent magnet is suspended between taut bands. It will attempt to line up with any external field, and an attached pointer and scale can be calibrated in kilogausses. Such a device can be made to indicate both direction and magnitude of the external field to a somewhat limited accuracy.

DC magnetic materials testing is done either by providing a complete closed path of the sample material on which exciting and sensing windings can be placed or by utilizing a "yoke" type of apparatus to furnish excitation to a small sample with its own sensing winding. Closed-loop samples may be a toroid composed of a stack of punched rings, a toroid made by wrapping tape into a spiral, or a closed loop made by stacks of strip samples assembled with overlapped ends in an Epstein frame.

This arrangement, in the form of a square, has an excitation winding and a sensing winding distributed along the four sides of the square to enclose the sample. The geometry and construction of these coils is detailed in ASTM Standard A343, part 44 of the *Annual Book of ASTM Standards*. Punched-ring samples are not usually considered satisfactory for oriented materials, while either spiral-wrapped tape toroids or Epstein strip samples can be used in either oriented or nonoriented materials. In any of these closed-loop samples, the excitation can be determined in terms of the ampere-turns that supply it. If the mean diameter of the sample is large compared with its radial width, the excitation is calculated as $H = 0.4\pi NI/l$ oersteds, where N is the number of turns in the magnetizing winding, I is the current in amperes, and l is the mean path length of the ring in centimeters. In using Epstein samples, it is necessary to make an assumption as to the actual magnetic-path length. This is normally taken as 94 cm in the 25-cm Epstein frame. A mutual inductor is included for calibrating the ballistic galvanometer, the series and parallel resistors in the galvanometer circuit permit adjustment to make the system direct-reading in appropriate units while preserving a desired galvanometer damping; resistors in the excitation circuit permit reversal or step changes at a desired ampere-turn level. Both excitation and test windings on the sample should be uniformly distributed.

Permeameters are used for small samples and for "hard" magnetic materials which cannot be driven to a sufficiently high excitation by readily applied turns on closed-loop specimens. Basically, all types of permeameters utilize heavy coils and large-cross-section yokes to provide a high excitation level in small samples.

There is increasing use of complete plotting systems for drawing magnetization curves and dc hysteresis loops. Such systems use a magnet assembly with tapered pole pieces adjustable with a screw drive for excitation of the sample.

Magnetic susceptibility testing designates those measurements which require much more sensitive apparatus than the methods described above. Such tests are made by measuring the minute mechanical forces experienced when the sample is in a nonuniform field. All these systems—the Gouy, the Faraday, and the Thorpe-Senftle method—consist of a strong field in which the sample is placed and weighed. They differ in the method of obtaining a calculable nonuniform field.

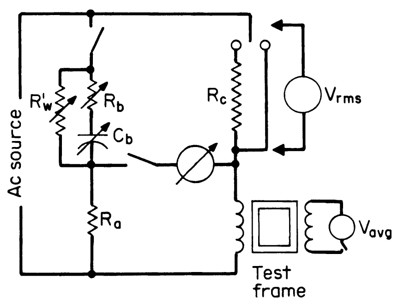


FIGURE 3-40 Modified Hay bridge.

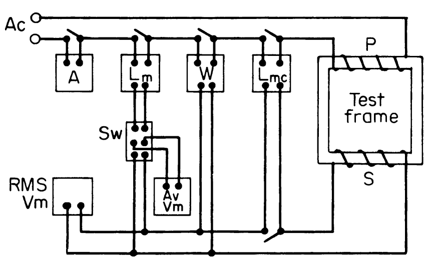


FIGURE 3-41 Voltmeter-wattmeter core-loss test system.

AC magnetic materials testing consists commercially in the determination of ac permeability and core loss in sheet materials. Substantially, all such testing is done either in Epstein-frame samples or in EI-type laminations. Up to an induction of 6000 G, measurements are made with the modified Hay bridge of Fig. 3-40. Above this level, measurements are made by the voltmeter-wattmeter method; Fig. 3-41 shows the circuit of such a test system. *A* is an ammeter of low impedance, *W* is a wattmeter with low-current circuit impedance and designed for low-power-factor use, rms V_m and av V_m are, respectively, rms responding and average responding voltmeters of very high impedance, L_m is a mutual inductance used with av V_m to read I_{peak} currents, and L_{mc} is a mutual inductance to compensate for the empty-frame mutual inductance of the Epstein frame. In operation, the flux density *B* is set using the average-responding voltmeter and calculating from the equation $4.444ANfB_{\max}/10^8 = 1.11\ E_{\text{av}}$, where $B_{\max}$ is the maximum induction in gauss, *A* is the cross section of the sample, *N* is the number of turns in the secondary (700 for the standard Epstein frame), and *f* is frequency in hertz. The value of *H* is determined by the formula $0.4NI_{\text{peak}}/L = H$ oersteds, where *N* is the number of turns in the magnetizing winding (700 for Epstein frame), I_{peak} is the peak current in amperes (derived from the reading of the voltmeter on the secondary of L_m), and *L* is the magnetic path length (94 cm for the 25-cm Epstein frame).

Core loss is calculated from the wattmeter reading divided by the active weight of the sample. Cross section is determined by weight of sample rather than an actual measurement of lamination thickness, with corrections for density of the material and assumed path length. Voltmeter-wattmeter measurements of core loss and ac permeability ($B_{\max}/H_{\text{peak}}$) are made with the actual instruments in the simple system. Commercial units for high-level production follow the basic circuit and include computation circuits to provide readings directly in the desired units, with printout of the data optional.

Magnetic amplifier material testing is a specialized procedure for materials to be used in amplifiers. There are a number of special tests in use on a supplier-user agreement basis that have no universal acceptance. ASTM Bulletin A598-69 specifies a number of recommended test points for various materials. These tests have the largest acceptance of any presently in use, and most suppliers are equipped to furnish material based on this type of testing. Test frequencies most commonly used are 60, 400, and 1600 Hz.

3.2 MECHANICAL POWER MEASUREMENTS

3.2.1 Torque Measurements

Torque is best measured with dynamometers, of which there are two classes: absorption and transmission. Absorption dynamometers absorb the total power delivered by the machine being tested, whereas transmission dynamometers absorb only that part represented by friction in the dynamometer itself. Made in a wide variety of forms, typical forms are described in the following paragraphs.

The Prony brake is the most common type of *absorption dynamometer*. The torque developed by the machine to overcome the friction is determined from the product of force required to prevent rotation of the brake and the lever arm. The load is applied by tightening the brake band or adding weights.

The *energy dissipated* in the brake appears in the form of heat. In small brakes, natural cooling is sufficient, but in large brakes, special provisions have to be made to dissipate the heat. Water cooling is the usual method, one common scheme employing a pulley with flanges at the edges of the rim which project inward. Water from a hose is played on the inside surface of the pulley and collected again by means of a suitable scooping arrangement. About 100 in^2 of rubbing surface of brake should be allowed with air cooling or about 25 to 50 in^2 with water cooling per horsepower.

The *Westinghouse turbine brake* employs the principle of the water turbine and is capable of absorbing several thousand horsepower at very high speeds.

In the *magnetic brake,* a metallic disk on the shaft of the machine being tested is rotated between the poles of magnets mounted on a yoke which is free to move. The pull due to the eddy currents induced in the disk is measured in the usual manner by counteracting the tendency of the yoke to revolve. This form of brake can be made in very small sizes and is therefore convenient for very small motors.

The principal forms of *transmission dynamometers* are the torsion and the cradle types. In *torsion dynamometers,* the deflection of a shaft or spiral spring, which mechanically connects the driving and driven machines, is used to measure the torque. The spring or shaft can be calibrated statically by noting the angular twist corresponding to a known weight at the end of a known lever arm perpendicular to the axis. When in use, the angle can be measured by various electrical and optical methods. The *cradle dynamometer* is a convenient and accurate device which is extensively used for routine measurements of the order of 100 hp or less. An electric generator is mounted on a “cradle” supported on trunnions and mechanically connected to the machine being tested. The pull exerted between the armature and field tends to rotate the field. This torque is counterbalanced and measured with weights moved along an arm in the usual manner.

3.2.2 Speed Measurements

Tachometers, or *speed indicators,* indicate the speed directly and thus include the time element. The principal types are centrifugal, liquid, reed, and electrical. In the *centrifugal type,* a revolving weight on the end of a lever moves under the action of centrifugal force in proportion to the speed, as in a flyball governor. This movement is indicated by a pointer which moves over a graduated scale. In the *portable* or *hand type,* the tachometer shaft is held in contact with the end of the shaft being measured, and in the *stationary type,* the instrument is either geared or belted. In the *liquid tachometer* of the Veeder type, a small centrifugal pump is driven by a belt consisting of a light cord or string.

This pump discharges a colored liquid into a vertical tube, the height of the column being a measure of the speed.

Reed tachometers are similar to reed-type frequency indicators, the reeds being set in resonant vibration corresponding to the speed of the machine. The instrument may be set on the bed frame of the machine, where any slight vibration due to the unbalancing of the reciprocating or evolving member will set the corresponding reed in vibration. Some forms are belted to the revolving shaft and the vibrations imparted by a mechanical device. *Electrical tachometers* may be either reed instruments operated electrically from small alternators geared or belted to the machine being measured or ordinary voltmeters connected to small permanent-magnet dc generators driven by the machine being tested.

Chronographs are speed-recording instruments in which a graphic record of speed is made. In the usual forms, the record paper is placed on the surface of a drum which is driven at a certain definite and exact speed by clockwork or weights, combined with a speed-control device so that 1 in on the paper represents a definite time. The pens which make the record are attached to the armature of electromagnets. With the pens in contact with the paper and making a straight line, an impulse of current causes the pen to make a slight lateral motion and, therefore, a sharp indication in the record. This impulse can be sent automatically by a suitable contact mechanism on the shaft of the machine or by a key operated by hand. The time per revolution is then determined directly from the distance between marks.

Stroboscopic methods are especially suitable for measuring the speed of small-power rotating machines where even the small power required to drive an ordinary speed counter or tachometer would change the speed, also for determining the speed of machine parts which are not readily accessible or where it is not practicable to use mechanical methods or where the speed is variable.

One convenient form of stroboscopic tachometer employs a neon lamp connected to an oscillating circuit supplied from a 60-Hz circuit, which is adjusted to “flash” the neon lamp at the frequency necessary to make the moving part that the lamp illuminates appear to stand still. Speeds from a few hundred to many thousands of revolutions per minute can be very conveniently measured.

3.3 TEMPERATURE MEASUREMENT

Temperature Scale. There is an international temperature scale, ITS-90 (International Temperature Scale of 1990), that came into effect on January 1, 1990 and superseded the IPTS-68 (International Practical Temperature Scale of 1968). All temperature measurements should be referred to the ITS-90. This scale extends upward from 0.65 K. The scale is defined in terms of ^{3}He and ^{4}He vapor pressure versus temperature relations, an interpolating constant-volume gas thermometer that is calibrated according to a specified procedure at designated fixed points to which temperature values have been assigned and that is used for interpolation according to specified equations, a set of defining fixed points to which temperature values have been assigned, and platinum resistance thermometers that are calibrated at specified sets of those fixed point and used for interpolation between those points according to designated reference and deviation functions, and Planck’s radiation law referenced to any one of three fixed points to which temperature values have been assigned. Temperatures on the ITS-90 were in as close agreement with Kelvin thermodynamic temperatures as possible at the time the scale was adopted. The scale is maintained in the United States by the National Institute of Standards and Technology (NIST), and any laboratory may obtain calibrations from NIST based on this scale. In the region from 0.65 to 25 K, rhodium-iron resistance thermometers and/or germanium resistance thermometers are calibrated on the ITS-90 to an uncertainty of $\pm$0.001 K or less; in the range from 13.8 to 934 K, standard platinum resistance thermometers (SPRTs) are calibrated to an uncertainty of $\pm$0.001 K or less; in the range from 273 to 1235 K, high-temperature SPRTs (HTSPRTs) are calibrated to an uncertainty of $\pm$0.002 K. Above 1235 K, the ITS-90 is realized by means of Planck’s radiation law, usually by calibrating a radiation thermometer, or pyrometer, at the freezing-point temperature of silver (1235 K), gold (1337 K), or copper (1358 K) and extrapolating to higher temperatures. The accuracy of the ITS-90 and the procedures used to calibrate the thermometers just described may be found in the references in the Bibliography. The remainder of this section on thermometry will be devoted to thermometry at a less accurate but more practical level.

Thermoelectric Thermometers (Thermocouples). By far the most commonly used thermometer in practical situations is the thermocouple. It consists of a pair of dissimilar electrical conductors (usually wires) joined at two junctions. One junction is maintained at a reference temperature t_0 (usually the melting point of ice), while the other is maintained at the unknown temperature t. The temperature difference produces a thermal emf which is measured by a potentiometer or a precise digital voltmeter. The latter is especially appealing because it is automatic (i.e., self-balancing), of sufficient resolution, and may easily be interfaced to an automatic data acquisition system.

Metals Used for Thermocouples. There are eight combinations of metal and alloys most extensively used, and they are designated as type B, E, J, K, N, R, S and T. Table 3-2 gives their nominal composition, temperature range, and highest suitable temperature for short-term use without significant degradation in performance.

Types R and S may be used for temperatures up to 1480°C and type B to 1700°C. It is recommended that the wire diameters exceed 0.5 mm if the thermocouple is to be used for long times at the upper temperature. These thermocouples are recommended for use in air because they are made from noble metals which are resistant to oxidation. They are easily degraded by other conditions, however, so they should be enclosed in a protective sheath. Type J may be used in a vacuum, inert, oxidizing, or reducing atmosphere. Again, a large-diameter wire (at least 3 mm) is necessary for use at long times in an oxidizing atmosphere. Types K and N are used up to 1200°C in inert or oxidizing atmospheres. Type E thermocouples are especially suitable for cryogenic use and may be used in vacuums, inert, oxidizing, or reducing atmospheres. Type T thermocouples may be used in the same atmospheres as type E, but they should not be used above 370°C under oxidizing conditions.

Temperature-EMF Relations for Various Thermocouples. Standard emf versus temperature tables, based on the ITS-90, have been developed and are published for the standardized thermocouples in NIST Monograph 175. Most manufacturers produce wires of sufficient quality so that a thermocouple may be fabricated from the materials given in Table 3-2, and their emf-t relation will deviate only slightly from that given in NIST monograph 175.

It must be understood that performance will degrade with use. There are a number of factors which cause decalibration, such as the atmosphere to which they are held at temperature and the highest temperature used. These effects are discussed in detail in the Bibliography.

Reference Junction Corrections. The values cited are appropriate for the situation in which the reference junction is maintained at the ice point ($t_0 = 0$°C). If the reference junction is not

TABLE 3-2 Standardized Thermocouples

Type designation	Nominal composition	Range, °C	Highest t for short-term service, °C
Type B	Pt − 30% Rh vs. Pt − 6%% Rh	0 to 1820	1700
Type E	Ni − 10% Cr vs. Cu − Ni*	−270 to 1000	370 to 870†
Type J	Fe vs. Cu − Ni*	−210 to 1200	320 to 760†
Type K	Ni − 10% Cr vs. Ni − Al	−270 to 1370	760 to 1260†
Type N	Ni − 14% Cr − 1.5% Si vs. Ni − 4.5% Si − 0.1% Mg	−270 to 1300	760 to 1260†
Type R	Pt − 13% Rh vs. Pt	−50 to 1768	1480
Type S	Pt − 10% Rh vs. Pt	−50 to 1768	1480
Type T	Cu vs. Cu − Ni*	−270 to 400	150 to 370†

*These alloys contain roughly 55% Cu and 45% Ni, and they are known as constantan.
†The highest temperature depends on the diameter of the wire. See ASTM Standard E230, Table 2, for further explanation.

maintained at that temperature, an emf correction must be applied to the measured emf to account for this. Refer to ANSI-MC96.1 for a discussion of this correction.

Extension (or Compensating) Wires. In many situations, it may be necessary for the reference junctions to be very distant (as much as several hundred feet) from the junction measuring T. Since most of the total emf in the thermocouple is generated by the short section at elevated temperatures, very little measurement error will occur if the remaining length at room temperature is replaced by thermocouple “extension” wires. Extension wires are made from materials having nearly the same thermal emf properties as the original thermocouple but which can be made at considerably less cost. For types E, J, K, N, and T thermocouples, the extension wires are made from the same alloy as the thermocouple wire but with less stringent requirements for the composition. For types R and S thermocouples, copper wire is used for one arm of the thermocouple, while a Cu–Ni alloy wire is used for the other.

Thermocouples: Summary. Thermocouples are relatively inexpensive, small, have rapid thermal response, and produce a signal (i.e., the emf) which is easily measured by a digital voltmeter. Spools of the thermocouple wire may be purchased, and many thermocouples may be made from them. Furthermore, each thermocouple will have an emf versus temperature given to within specified tolerance of the standard table so that in many instances calibration is not necessary. There are disadvantages to thermocouples, however; the emf is sensitive to the temperature distribution along the wire, strains, thermal history, and degradation at elevated temperatures. If these latter problems outweigh the advantages, other thermometers described below may be more appropriate.

Liquid-in-Glass Thermometry. A liquid-in-glass (LIG) thermometer consists of a reservoir of liquid and a stem with a temperature scale marked on it. The liquid has a very large thermal expansion compared with the reservoir and stem, and thus small temperature changes cause the liquid to expand into the stem where the length indicates T. A wide variety of liquids and glasses are used, but the most common liquid is mercury enclosed in borosilicate glass. If properly treated, LIG thermometers are capable of repeatedly measuring T to within 0.03°C. The major cause of catastrophic failure is breakage; of noncatastrophic failure, heating the thermometer beyond its specified range. LIG thermometers are used widely throughout industry because they are inexpensive and easy to read with the human eye. They are not amenable to automation or continuous monitoring, however. In many applications, they are being replaced by thermocouples or resistance thermometers, commonly called *resistance temperature detectors* (RTDs).

Resistance Temperature Detectors. Since resistance is a physical property that is easy to measure and automate with modern instrumentation, RTDs are finding more general acceptance in temperature measurement. The two major classes of resistors with strong temperature dependence are thermistors and platinum resistors.

Thermistors. Thermistors are made by sintering mixtures of oxides of Mn, Fe, Co, Cu, Mg, or Ti, bonding two electrical leads to the sintered material, and enclosing the unit in a protective coating. The devices are made in a wide variety of shapes (beads, disks, rods, and flakes), are very inexpensive, and are very compact (one bead type commercially available is only 0.07 mm in diameter). The resistance of the device is generally high (1 to 100 kΩ), so lead resistance is not a significant source of measurement error. If glass-encapsulated bead-type thermistors are used below 150°C, they are quite stable. (Commercial units are available which drift by no more than 0.01°C per year.)

Thermistors are semiconducting devices whose resistance depends exponentially on temperature. This means that the thermistor is very sensitive to temperature, but it also means that its temperature range is limited (i.e., if T becomes too low, the resistance becomes too high to measure; if T is too high, the resistance becomes too low to measure). Thermistors may be chosen for use with temperatures as low as 4 K, while others may be used in the region near 900 K. A technique, referred to as *linearization,* may be used to extend the operating range of a thermistor. This consists of connecting a temperature-independent resistor R in parallel with the thermistor. If the value of the resistor is equal to that of the thermistor in the center of its operating range, the resistance of the circuit will be roughly linear in temperature rather than exponential.

Thermistor-based measurement systems with digital readouts which read directly in temperature are widely available. These consist of a sensor, a digital ohmmeter, and a logic unit, generally a microprocessor. The microprocessor is used to perform resistance-to-temperature conversion and perhaps integration, control timing, run a display, and provide digital output for computer analysis.

Platinum Resistance Thermometers (PRTs). The resistance of platinum is roughly linear in temperature over a very wide temperature range, and thus PRTs may be used over a greater temperature range than thermistors. Precision-type PRTs can be very reproducible and are capable of high accuracy. They are, however, more sensitive to mechanical shock and less sensitive to temperature change than are thermistors.

For the highest-accuracy temperature measurements, three types of "standard" PRTs are used. From 83.8 to 693 or 934 K, a well-characterized, fine Pt wire is supported by insulators and enclosed in a fused silica glass casing. The assembled unit is 600 mm long and 7 mm in diameter. From 13 to 84 or 273 K, a "capsule" version 60 mm long, 6 mm in diameter, with similar internal construction, is used. A third type, called a *high-temperature standard PRT* (HTSPRT), is constructed of larger wire and typically has a resistance of about 0.25 or 2.5Ω at 273 K. That wire is wound on fused silica formers, and the overall length of the unit is about 24 in. When properly used (see NIST Technical Note 1265), PRTs may be used to measure temperature with an imprecision not exceeding ±0.001 K over the range 13 to 934 K. Such standard PRTs are used to realize the ITS-90 over the range 13 to 1235 K. Since great care must be exercised in measuring and handling these devices in order to achieve this performance, standard PRTs are generally restricted to the primary standards laboratory of any organization.

Thermal Radiation. Radiant flux, in the form of photons or electromagnetic waves, emitted by a surface solely as a consequence of its temperature is known as *thermal radiation*. Its wavelength range extends from about 100 nm (far ultraviolet) through the visible to about 1 mm (far infrared).

Blackbody Radiation. A surface which absorbs all incident radiation, regardless of wavelength or direction, is known as a *blackbody*. For a given temperature and wavelength, such a surface will emit the maximum possible thermal radiation. The thermal power emitted by an element of area into an element of solid angle surrounding the given direction of propagation (the *radiance*) is independent of direction for a blackbody. The spectral radiance distribution of a blackbody, $L_b(\lambda, T)$, is described by the Planck equation:

$$L_b(\lambda, T) = \frac{c_1}{\lambda^5(e^{c2/\lambda T} - 1)}$$

where λ is the wavelength of the emitted flux, T is the thermodynamic temperature of the blackbody, c_1 is the first radiation constant (3.741832×10^{-16} W · m^2), and c_2 is the second radiation constant (1.438786×10^{-2} m · K). A blackbody does not exist, but it can be closely approximated by a very small opening in a uniformly heated hollow opaque enclosure.

Stefan-Boltzmann Law. The relation between the total power per unit area emitted by a blackbody M_b and its thermodynamic temperature T is expressed by the equation

$$M_b = \sigma T^4$$

where σ is the Stefan-Boltzmann constant (5.67032×10^{-8} W · m^{-2} · K^{-4}). This equation is obtained by integration of the Planck equation over the wavelength range zero to infinity.

A real surface can never emit more thermal power than a blackbody. It is often convenient to express the total thermal power emitted by a real surface per unit area, M_s, to that emitted by a blackbody at the same temperature by the equation

$$M_s = \epsilon M_b = \epsilon \sigma T^4$$

where ϵ is the total hemispherical emissivity (note that $\epsilon < 1$).

Radiation Thermometer (Pyrometer), Wide-Band. In a wide-band radiation thermometer, thermal radiation from the source is focused on the sensor by means of a lens or concave mirror. The sensor might typically be a thermopile, a pyroelectric, or a solid-state photodiode detector. The fraction of the total thermal radiation received from the source is limited by the spectral transmission/reflection characteristics of the components in the optical path as well as the spectral response of the sensor itself. For measurement of the temperature of a blackbody source, the relation between the output of the sensor S and the temperature of source T_s can be approximated by the equation

$$S = a\sigma T_s^b$$

where a and b are instrumental constants obtained through calibration. With suitable linearization circuitry, the output of the instrument can indicate temperature.

Radiation Thermometer (Pyrometer), Spectral-Band. In a spectral-band radiation thermometer, thermal radiation from the source is typically focused on the sensor by means of a lens. The bandwidth of the thermal radiation reaching the sensor is defined by a transmission filter placed in the optical path within the instrument. The sensor found in most instruments is a solid-state photodiode detector. For measurement of the temperature of a blackbody source, the output of the detector is approximately proportional to the Planck equation.

With suitable linearization circuitry, the output of the instrument can indicate temperature. Many narrow-band instruments have a control which can be adjusted by the operator to compensate for the emissivity of the source.

3.4 ELECTRICAL MEASUREMENT OF NONELECTRICAL QUANTITIES

A *transducer* is a device in which variations in energy of one form produce corresponding variations in energy of another form. In common usage, either the input or output of a transducer is electrical. Thermocouples and thermistors fall into that category, as does the *thermal converter,* whose electrical output (dc millivolts) is derived from a thermal effect that represents an electrical quantity (ac volts, current, watts, vars) that differs in nature from the output. A variety of methods is often available for the measurement of a specific variable. "Frequently, operational considerations will indicate the choice of transducer; for instance, piezoelectric transducers may not perform well if long cables are required; capacitive devices, although quite sensitive, may require intermediate electronic circuitry; and magnetic transducers should not be used in the presence of strong magnetic fields."

Mechanical displacement may be converted into an electrical variable by the simple expedient of adjusting resistance in an electric circuit. A slide-wire resistor, having a movable contact attached to the part whose displacement is to be measured, may be connected through a 2-conductor circuit to a steady-voltage source in series with an ammeter (or milliammeter) calibrated in terms of the displacement. If the resistor is connected as a voltage divider, the need for a regulated supply is eliminated, and with a 3-conductor circuit the display instrument may be a ratio meter or a potentiometer. Such combinations are common and are available for both dc and ac operation. Where deflections are small—less than 0.1 in—measurement may be made by use of a differential transformer.

In the *strain gage,* microscopic relative displacements are electrically magnified and are displayed on an indicating or a recording meter or on an oscillograph. Modern resistance-type strain gages comprise fine-wire windings arranged to be more or less elongated when subjected to deformation. The units may be used singly, in pairs, or in sets of four constituting a complete Wheatstone bridge. There are two main classes of wire-wound strain gages, (1) bonded and (2) unbonded.

1. The *bonded strain gage* is composed of fine wire, wound and cemented on a resilient insulating support, usually a wafer unit. Such units may be mounted on or incorporated in mechanical elements or structures whose deformations under stress are to be determined. While there are no limits to the basic values which may be selected for strain-gage resistances, a typical example may be taken as of the order of 100 to 500 Ω.

2. In the *unbonded strain gage,* the resistance structure comprises a fine-wire winding stretched between insulating supports mounted alternately on the two members between which displacement is to be measured (Fig. 3-42). These wires comprise the four arms of a Wheatstone-bridge network of which two opposite arms are tightened and the other two are slackened by the displacement. While a bonded gage tends to respond to the average strain in the surface to which it is cemented, the unbonded form measures displacement between the two points to which the respective supports are attached. Unbonded wire strain gages are usually operated on input potentials ranging up to 35 V direct or alternating current. Under conditions of extreme unbalance, corresponding to full operating range, the open-circuit emf may be of the order of 8 to 10 mV and the closed-circuit current up to 100 μA.

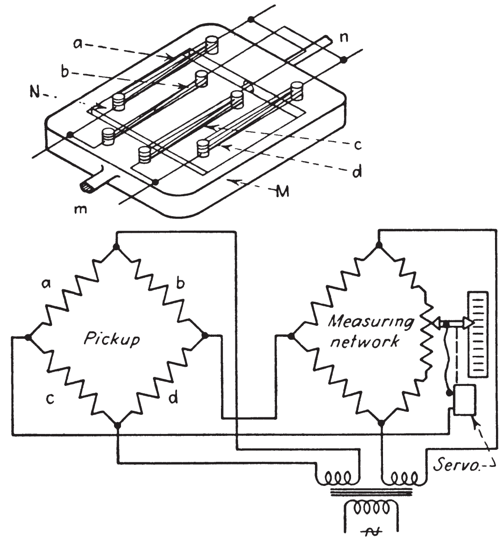


FIGURE 3-42 Diagram of unbonded wire strain gage. Supports *M* and *N* are attached by rods *m* and *n*, respectively, to points between which displacement is to be measured. Pickup and measurement networks are energized from similar but isolated sources. Unbalance originating in the pickup is detected and balanced by a servo-actuated measuring network, providing a reading of strain on a graduated scale.

Recently developed types of conductive rubber are used in resistive transducers capable of wider ranges of deformation than are those using wire or foil. Where the strain gage must operate over a temperature range, dummy gages exposed to the temperature but not the strain may be employed for temperature compensation, or alloys having a low temperature coefficient of resistance may be used. Piezoelectric strain gages are also available for applications in pressure, force, torque, and displacement measurement. Strain gages for use on ac circuits are supplied in both capacitive and inductive forms, wherein the corresponding characteristics of ac circuit components are varied by the displacements to be measured.

A popular means for measuring *small displacements* in the range from a millimeter to a micron is the *linear,* or *differential, transformer*. This device is generally produced with a single primary winding and two secondaries, all disposed along a common axis and having in the common magnetic circuit a movable iron core longitudinally displaceable with the motion to be measured. The secondaries may be connected additively or differentially and may be included in the circuit of a null-type instrument balanced either by shifting the core of a similar transformer excited from the same source or by the use of a slide-wire potentiometer. Linear transformers are regularly supplied for operation at all frequencies up to 30,000 Hz. The sensitivity, of course, increases with the frequency. Linear transformers may be interconnected in a great variety of arrangements to perform computations or to express desired mathematical functions of measured variables.

Strain gages permanently attached to diaphragms, tubes, and other pressure-sensitive elements find wide applications as components of *electrically actuated pressure gages*. By electrically combining simultaneous measurements of torque and velocity, continuous determination of mechanical power may be obtained, the combination becoming an electrical-transmission dynamometer.

Vibration may be determined by a strain gage, but the fact that this magnitude involves motion renders it generally preferable to utilize alternating potentials developed by periodic change in the geometry of the measuring circuit. This may be embodied in either a *capacitor* or an *inductive device*. In a recently developed apparatus, there are no moving parts except the object being shaken, and the vibration displacement is sensed by its effect on an electrostatic field between the pickup and the moving part. *Piezoelectric crystals* are particularly adapted to the measurement of vibration. The emf so obtained is proportional to the amplitude of deflection multiplied by the frequency squared.

Air velocities and the *flow of gases* in general may be measured by the *hot-wire anemometer*. In its simplest form, this device utilizes the cooling effect of the gas stream to establish a temperature difference between exposed and protected bridge arms. Where the flow is in an enclosed conduit, a heating element may be introduced and the volume of flow determined by the amount of heat transferred between the heater and the temperature-sensitive bridge wires.

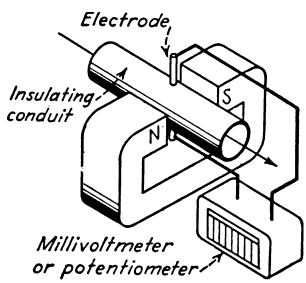


FIGURE 3-43 Electromagnetic flowmeter.

Flow of an electrically conducting liquid may be determined by measuring the emf developed between a pair of electrodes set in opposite sides of an insulating conduit due to the movement of the liquid through a magnetic field established transversely of the conduit and perpendicular to both the flow and the line joining the electrodes (Fig. 3-43). By using an alternating field, the effects of electrode polarization may be eliminated. Null measurement of the generated voltage renders the apparatus independent of the resistance of the liquid.

Liquid level may be expressed electrically by the use of a transducer responsive to the verticalposition of a float or by a pressure-sensitive strain gage immersed in the liquid below its lowest level. Variation in resistance of an immersed conductor is a widely accepted principle, especially in fuel tanks. If the liquid is an electrical insulator and of constant characteristics, its depth may be determined by its dielectric effect between a pair of vertically disposed capacitor plates. On the other hand, if the liquid is a conductor and very small changes in level are to be detected or regulated, the liquid may be made one electrode of a capacitor whose other electrode is a horizontal plate positioned above the surface.

Levels of corrosive liquids or those operating under extreme pressures, temperatures, or other conditions rendering them inaccessible for measurement by conventional means may be determined by the use of *gamma radiation*. Several gamma-ray sources are spaced at equal vertical intervals in the tank or reactor containing the liquid to be measured but are positioned so that none of them obstructs the line of sight of a Geiger counter tube placed at the top of the container. The response of the Geiger tube depends on the depth of the process material, and the output is measured on a null-type recording instrument.

Vacuum may be measured by determining either *energy dissipation* or electron emission in the space under test. The former principle provides the basis of the *Pirani gage*, wherein two similar heated filaments forming arms of a bridge are located, respectively, in a reference bulb and a bulb connected to the evacuated space. Heat dissipation will vary with the degree of evacuation, while conditions in the reference bulb remain constant. The electrical condition of the bridge then provides a continuous measure of the vacuum. The normal range of operation of the Pirani gage is from 10^{-7} to 5 torr. Since the performance of a thermionic tube is highly responsive to the degree of vacuum, its action under controlled electrical conditions is a criterion of internal atmosphere. This principle forms the basis of a number of *electronic vacuum gages*. The normal range of operation lies between 10^{-7} and 10^{-3} torr.

Electrical methods for *analyzing gases,* while essentially thermal in their nature, are made practicable only by the application of electrical principles in determining thermal relationships. In the *thermal-conductivity method,* as best exemplified in the CO_2 recorder, two cells or sections of conduit containing, respectively, a standard sample and the gas under test have in them adjacent arms of a bridge network composed of wires having known resistance variation with temperature and carrying sufficient current to raise their temperatures appreciably above their surroundings. As more or less heat is dissipated in the test cell as compared with the reference cell, the relative resistance of the bridge arms varies, providing an electrical basis for measurement of the gas composition.

The *catalytic-combustion method* is especially adapted to *detection of flammable gases* or determination of explosibility. The arrangement of cells and bridge wires may be similar to that of the thermal-conductivity type, but the filament is usually composed of activated platinum and is operated at a temperature sufficient to ignite the gas when a critical proportion is attained. The increased heating of the bridge wire due to combustion abruptly disturbs the balance and provides a positive indication of explosibility. In some forms of this instrument, the temperature rise is determined by thermocouples. The catalytic-combustion method is useful in determining mixtures containing such gases as propane, acetone vapor, carbon disulfide, and carbon monoxide. The equipment finds use in (1) solvent-recovery processes, (2) solvent-evaporating ovens, (3) combustible-gas storage rooms, (4) storage vaults, (5) gas-generating plants, (6) refineries, and (7) mines.

In determining the *oxygen content of gases,* both the conventional thermal-conductivity method and the catalytic combustion method are applicable. In addition to these, use is made of the *magnetic susceptibility* of oxygen as a basis of operation. In one such instrument, a hot-wire bridge similar to that of a CO_2 recorder is employed, one of the gas chambers being placed in a strong magnetic field. This stimulates the flow of oxygen-containing gas through that chamber, thereby unbalancing the bridge by a measurable amount. In the other magnetic analyzer, a test chamber contains a small magnetic member rotatable in a distorted field whose conformation depends upon the amount of oxygen present. The resultant angular displacement of the test member may be used similarly to that of a galvanometer in either a direct-deflecting or a null-type instrument.

An analyzer especially suited to measurement of *toxic ionizable gases* or vapors to and beyond the toxic limits utilizes the electrical conductivity of an aqueous solution of the gas. The vapor under test is bubbled through distilled water at a fixed rate, and the conductivity of the solution becomes a measure of gas concentration. A typical use is the continuous recording of small quantities of substances like sulfur dioxide, hydrogen sulfide, chlorine, and carbon disulfide in the air.

Atmospheric contamination may be determined by an *electronic leak detector,* utilizing emission of positive ions from an incandescent filament exposed to the air. The filament is enclosed in an open inner cylinder and heated by alternating current. The atmosphere under test is forced through the annular space between the inner and an outer cylinder at a predetermined rate, and the electron flow due to a dc potential maintained between the cylinders is measured as an index of the amount of contaminant. Presence of extremely small proportions of halogen vapor compounds, of which Freon, chloroform, and carbon tetrachloride are good examples, greatly increases the emission. At room temperatures, the device does not respond to Pyranol, but if this material is heated sufficiently to give off vapor, a response is obtained. It also responds to solid particles of the halogens and therefore will detect *smoke* from burning materials containing these elements. The instrument is also available as a recorder and/or a controller.

Relative humidity is determinable electrically by methods involving either of the two basic principles: (1) variation of electrical conductivity or of dielectric constant of a hydrophilic element and (2) computation based on "dry-bulb" and "wet-bulb" temperatures of the atmosphere whose moisture content is to be determined. The most common embodiment of the former method consists of an insulating card, plate, or cylinder carrying a bifilar winding of conductive wire and having a relatively large surface exposed to the atmosphere. The two strands of wire are bridged by a coating of material such as lithium chloride or colloidal graphite, having a high affinity for moisture. This material quickly assumes a water content corresponding to that of the atmosphere, and the electrical resistance between the conductors becomes a function of the humidity to be measured. A similar principle is used in determining the moisture content of hygroscopic materials such as wood, grain, or pulp. In such applications, a resistance-measuring circuit terminates in electrodes or probes which are pressed against or inserted into the material to be tested.

Moisture content of material in a web or sheet form, such as paper, may be continuously determined by passing the web between the plates of a capacitor, and thus obtaining a measurement determined by the dielectric constant of the material as affected by its water content.

Electrical determination of humidity by the wet-and-dry-bulb method requires somewhat intricate computing circuits which for accurate results must take account of absolute temperature and barometric pressure.

Determination of dew point, or the temperature at which condensation takes place on a polished surface, as a function of absolute humidity, employs essentially a thermal and optical method of measuring, but such a system may be rendered continuous and automatic by photoelectrically observing the conditions of a polished surface in the tested atmosphere utilizing a servo system to regulate its temperature and thus obtaining an indication or a record of the dew point.

The two most popular types of *electric micrometers* are (1) that utilizing the magnified output of a strain gage and (2) that based on precise determination of capacitance between two electrodes whose spacing corresponds to the measured dimension.

Ultrasonic thickness gages may be used to measure steel walls ranging in thickness from ⅛ in to 1 ft, utilizing the fact that sound vibrations tend to establish standing waves within the mass of the material upon which they are impressed. This device combines a variable-frequency oscillator with a piezoelectric crystal which is pressed against the wall to be tested. The circuit is tuned until the metal oscillates, causing a sharp increase in the loading. The frequency of this resonance indicates the thickness of the material.

Selection of a method for determining the *thickness of sheet material* in process will depend primarily on the inherent electrical conductivity of that material. If it is essentially a nonconductor, such as rubber, plastic, or paper, measurement may be continuously performed by passing the sheet or web between the plates of a capacitor. (In such measurements on hygroscopic materials, moisture content may become a dominating factor.)

Sheet thickness in sheet materials, whether conducting or insulating, may be measured by the *beta-ray gage.* In this device, a stream of beta rays passes through the sheet to a pickup head whose response is amplified and continuously recorded and, if desired, made the controlling influence in automatic regulation. Provision is made for the combined radiation source and pickup to traverse the strip of material and scan its whole width.

Thickness of coatings, such as varnish or lacquer, on conducting materials may be determined by a continuous measurement of capacitance between the base and a reference electrode, the coating being included as a dielectric. With a magnetic base, such measurement may be performed effectively by determining the effect of the coating on the gap in a magnetic circuit.

Surface roughness may be determined either on an absolute basis or by comparison with a "standard" surface. A common method involves passing a small stylus systematically over the surface, similarly to a phonograph needle, and measuring the resulting vibration. The stylus may be attached to a strain gage, piezoelectric crystal, or a magnetic pickup. The resulting alternating emf may be amplified and displayed on an oscillograph, or it may be rectified and measured with a millivoltmeter. A basis for quantitative determination of surface roughness is found in USAS B46.2.

An absolute method of determining roughness uses the electrical capacitance of the tested surface in contact with an electrolyte as compared with that of an ideal (mercury) surface. On the assumption that the capacitance varies as the surface area, the comparison provides a figure representing the ratio of the tested surface to one of perfect smoothness.

Transparency (or opacity) determination of materials and continuous monitoring of smoke density involve passing the substance to be examined through the path of a light beam directed upon a photocell. Uninterrupted measurement is made by means of a potentiometer or a bridge, according to the class of cell employed.

Viscosity measurement is essentially mechanical in its nature, and the application of electrical methods consists of determination of stress or displacement set up in the measuring apparatus owing to the characteristic of the fluid. One method involves measuring the electrical input to a small motor driving an impeller or stirrer in the fluid. Another method is based on electrical determination of the angle of lag (torque measurement) in a resilient mechanism through which an impeller is driven. A further method utilizes *magnetostriction* to produce longitudinal oscillations in a steel rod carrying a diaphragm immersed in the liquid. Determination of the electrical loading on the exciting circuit provides a measure of viscosity.

Electrical measurement has superseded many of the older methods of quantitative determination of *chemical magnitudes*. The two best-known methods are based, respectively, on the electrical *conductivity* of solutions and on the *voltaic effect* in specific cells. The basic principles of these measurements are wholly different, as are their applications. In the *conductivity* cell, every precaution must be observed to avoid electrolytic effects, the prime requisite being that the respective electrodes be of identical material. Even then, the passage of current or the application of the potential tends to produce internal polarization emf in the cell. This undesirable effect may be almost wholly eliminated by measuring electrolytic resistance with alternating current, and the highly sensitive ac detectors now available enable such tests to be made with precision. Outstanding among the uses of the resistance cell is determination of the *purity of water* for domestic and industrial purposes. Conductivity of water solutions usually increases in proportion to the amount of dissolved electrolytic material. Perfectly pure water has a specific resistivity of 18 to 20 million Ω/cm^3, but in practice, such values are virtually unobtainable. Only by careful distillation or deionization is it possible to obtain water of 400,000 to 800,000 specific Ω at a reference temperature of 20°C. Continuously operating water-conductivity recorders are supplied for use with commercial ac power supply, and a typical range is 100,000 specific Ω to infinity.

Electrolytic cells utilize measurement of emf developed between a standard combination of electrodes by the solution under test. Development of the principle has reached its highest refinement in the measurement of pH, or hydrogen-ion concentration, which is a criterion of the activity with which the solution will enter as an acid into a chemical reaction. The pH value is a logarithmic function of the emf developed with a given strength of the solution in a specified cell. For pure water, which is "neutral" in its reaction, lying midway between the acids and the bases, the pH value is 7. The pH measurement is essential in practically every industry involving any chemical process, as well as in waterworks, sewage systems, biological laboratories, and agricultural experiment stations.

3.5 TELEMETERING

Telemetering is measurement with the aid of intermediate means which permit the measurement to be interpreted at a distance from the primary detector. The distinctive feature of telemetering is the nature of the translating means, which includes provision for converting the measurand into a representative quantity of another kind that can be transmitted conveniently for measurement at a distance. The actual distance is irrelevant.

Electric telemetering is telemetering performed by deriving from the measurand or from an end device a quantitatively related separate electrical quantity or quantities as a translating means. A *measurand* is a physical quantity, property, or condition which is to be measured.

Telemetering has been practiced many years in the central-station industry and in the transmission and distribution of electric power but until lately only to a limited extent in the nonelectrical fields. With the phenomenal expansion of pipelines for gas and for oil, the need has vastly increased, and electric telemetering installations have become indispensable in the remote measurement, totalization, regulations, and dispatching of these utilities. Telemetering also has found wide application in extensive industrial plants, such as refineries, steel mills, and large chemical plants, and in these installations it often forms an essential part of remote regulating apparatus.

There has been a rapidly increasing use of telemetering in aircraft, meteorology, ordnance, and guided missiles. This has led to a sharp demarcation of telemetering philosophies and techniques into two classes, *mobile* and *stationary*. In the former, the apparatus is expected to operate for a very short period of time—often only a matter of seconds. The transmitting unit at least must be considered as expendable, and the combination is generally subject to an overall calibration for each isolated test in which it is used. Obviously, there can be no interconnecting physical circuit, and a radio link is an essential part of the system.

Stationary systems in general involve transmitting and receiving units at fixed locations. These are usually of a permanent nature and are intended for operation over extended periods of time. Signal transmission between the stations usually involves a physical circuit, and even where radio principles are utilized, the most common practices require guiding of the signal by means of a more or less continuous conducting path.

A *telemetering system* incorporates the same three essential elements as are required in a system for measurement of nonelectrical quantities by electrical means, namely, a *transmitting unit* (transducer or pickup), a *receiving unit* (an instrument for measuring an electrical variable), and an *interconnecting circuit* or channel by which the electrical variable (signal) originating at the transmitter is carried to and impressed upon the receiver.

In transmission of measurement over considerable distances, the *circuit* or *channel* may become the predominating factor in the system. In the ideal telemetering system, the terminal apparatus would be inherently self-compensating so that variations in circuit conditions would not adversely modify the signal. Merit of a telemetering system is directly related to the degree to which it approaches this ideal. Distance criterion of a telemetering system is not so much the number of feet or miles over which it will operate as it is the nature and magnitude of circuit impedance through which its signals will maintain their identity and proportionality. Since the data have been determined for specific types of circuits and channels, such magnitudes generally may be expressed in units of distance. A continually increasing proportion of telemetering is being carried out over circuits and channels leased from communication companies. With information available respecting the type of signal to be transmitted, the telephone or telegraph company provides a suitable circuit and assumes responsibiltiy for its operation. Where privately owned circuits are used for telemetering, their maintenance and protection correspond to those for comparable communication circuits.

In *classifying telemetering systems,* the ANSI has adopted a grouping recommended by the AIEE and based on the nature of the electrical variable transmitted through the interconnecting circuit or channel. The names of the five classes are more or less self-explanatory and are as follows: *current, voltage, frequency, position,* and *impulse* types. In each of the first three of these classes, the corresponding characteristic of the electrical output of the transducer comprising the transmitting unit is varied with variations in the measurand. In the *position* system, the quantitative ratio, or the phase relationship, between two electric voltages or currents determines the nature of the transmitted signal, usually requiring a circuit of three or more conductors. There are several *impulse* systems, in all of which the transmitting instrument acts to "key" a signal impressed on the circuit, producing a series of successive pulses which, according to their nature, are interpreted by the receiving instrument and expressed in terms of the measurand.

Telemetering systems are not always mutually exclusive. A single installation may represent a combination of several of the named systems. In some instances, it becomes difficult to decide into which of the specified classes a particular method of telemetering may fall.

Telemetering of electrical quantities, such as volts, watts, and vars (volt-amperes reactive power) presents a problem owing to the inherently low torque of direct-deflecting instruments, whereas devices for measuring such quantities as position, flow, and liquid level are not subject to such restrictions. Accordingly, where measurements of electric units are to be transmitted, practice favors those systems which place a minimum of burden on the primary measuring instruments and preferably those adapted to transmitters having no moving parts. Thus, photoelectric, thermoelectric, and capacitive transmitters have found considerable favor in the electric industry.

In transmitting measurements originating in *integrating meters,* such as watthour or varhour meters, the mechanism of the meter, either by photoelectric or electronic means or by a contact arrangement, is caused to develop a series of electrical pulses whose frequency of occurrence is proportional to the instantaneous value of the measured load. By a simple electronic network including capacitors charged and discharged at the frequency of the pulses, there is produced a direct current whose value is proportional to that frequency, the telemetering system being thus placed in the *current* class. On the other hand, the pulses may be directly impressed on the communication channel, whereupon the system falls into the *frequency* group.

Where the basic measurement is performed by a low-torque instrument of the direct-deflecting class, such as a wattmeter, common telemetering practice involves either *balancing the torque* or matching the deflection of the instrument by the effect of an automatically regulated direct current in the winding of a permanent-magnet moving-coil mechanism. This current, remaining proportional to the instrument torque, is transmitted through a metallic circuit for measurement at the receiving station and, if desired, may be included with other and similar currents in a load totalization.

A most flexible method for the transmission and totalization of electric power measurements involves the use of a *thermal converter*. The several commercial forms of this device operate on a long-known but only recently applied principle combining the circuit of the thermal wattmeter with that of the thermocouple. In the former, the temperatures of two resistors are caused to assume values differing by an amount proportional to the power in the measured circuit. In the latter, there is developed an emf proportional to the temperature difference or to the *power* in the measured circuit, irrespective of power factor, frequency, or waveform. Thermal converters are supplied in single-element, two-element, and three-element forms, and the ac input circuits may be wired into the instrument-transformer secondaries on any conventional polyphase power system. The output from the dc terminals is either measured directly or interconnected with that of other converters to provide totals of measured loads. The full-load potentials are usually rated at 50 or 100 mV, according to make and type, and measurement is preferably made with a self-balancing potentiometer. For best results, thermal-converter output circuits, which, of course, must be wholly metallic, should be well shielded from parasitic electrical effects and preferably should be in a sheathed cable. An advantage of thermal-converter installations, even for relatively short distances within the plant, is that the seven or eight conductors necessary for connecting instrument-transformer secondaries to wattmeters or varmeters are replaced by two small wires operating at a negligible power level. Furthermore, physical damage to the output wiring, whether in the nature of an open circuit or a short circuit, is not hazardous to equipment or personnel, and on restoration of the circuit, normal operation will be resumed without loss of accuracy.

Electrical impulses may be used as signals for telemetering in a number of ways, the most important in stationary installations being that based on *frequency* and that based on *duration* of successive impulses. Impulse systems are to telemetering what telegraphy is to other forms of communication. The function of the transmitting instrument is essentially one of “keying” a circuit. Since the significance of the transmitted signal is based on time only, it follows that the method is most nearly immune to circuit conditions, such as voltage variation, impedance changes, attenuation, poor connections, and pickup from adjacent disturbing influences. Impulses whose frequency represents the measured variable may be transmitted as such, then falling into the category of the *frequency system of telemetering,* or they may be converted into a proportional direct current and be classified with the *current* systems.

Impulse-Duration Telemetering. Signals recur at uniform intervals, and each has a duration corresponding to the then existing value of the measured magnitude. The transmitting instrument includes a constantly running cam or scroll plate having a spiral trailing edge and operating in the plane of the pointer but perpendicular to the line of excursion. At a fixed point in each revolution of the cam, the pointer is engaged and brought against the cam face until subsequently released by the trailing edge. With engagement and disengagement, the pointer is slightly deflected perpendicular to its line of travel and actuates a contact in a signal circuit. Because of the spiral form of the trailing edge, the length of the signal depends on the position of the pointer and thus represents the measured variable.

The receiving unit includes the equivalent of a pair of electromagnetic clutches continuously driven by a constant-speed motor. These clutches are actuated by the incoming signals, one in an “upscale” and the other in a “downscale” sense, according to whether the transmitter pointer is on or off the cam. The receiver pointer or pen is frictionally retained in position and is “nudged” alternately toward one end or the other of its range by impellers or “dogs” carried by the clutches and, respectively, reset to zero as the corresponding clutch is released. Thus, with each signal, the receiver pointer finds or maintains a position corresponding to that of the transmitter pointer.

Position Telemetering. In the *position system* of telemetering, the characteristic signal involves the relationship between two electrical quantities of a similar nature, that is, two voltages or two currents. Unless carrier is used, position systems (with one exception) require an interconnecting circuit of three or more conductors. The simplest position-telemetering arrangements are those of the rheostatic, or bridge, type, either direct or alternating current. Mechanical attachment of the measuring element to a voltage-dividing resistor provides a transmitting unit wherein the relative value of two voltages may be made proportional to the measured quantity. The receiving instrument may take the form of a ratio meter or may be a self-balancing bridge. The accuracy of such systems is affected by the impedance of the interconnecting circuit, but by maintaining this value small in comparison with that of the terminal instruments, the error may be made negligible for considerable lengths of line.

Selsyn. The inductive type of position system is best exemplified in the “selsyn” position motor or any one of its several equivalents. The transmitting and receiving units may be identical in structure. Each involves a stator and a rotor, one being provided with a single-phase and the other with a polyphase winding. The single-phase windings are excited from a common ac source, and the polyphase windings are interconnected. The rotors of the two units will tend to assume duplicate angular positions so that if one is attached to a measuring element, the other will provide a remote indication of its position. This system requires three line conductors in addition to the pair comprising the common power supply. The versatility and flexibility of the *differential transformer* render it particularly adaptable to telemetering of mechanical displacements.

Totalization of power loads and of other measured quantities may readily be effected in the current or voltage systems by connecting the outputs of the respective transmitters in parallel or in series, as the case may be. Subtotals and other mathematical functions also may be obtained. Telemetering, especially totalization and retransmission, is greatly facilitated by the power and flexibility of servo-actuated potentiometers and bridges. With these instruments available, there is practically no limit to the possibilities of telemetering, not only in the electrical-utility field but, also in association with pipelines and large industrial plants. By *multiplexing* the circuits, it is possible for several telemetering transmitters and receivers to share a common communication channel. The most common systems of multiplexing are those based on *frequency* and those based on *time*. The frequency method transmits the signals on carriers having a specific frequency allotted to each transmitter and receiver combination. Time multiplexing involves the use of a multiple-point switch at each end of the circuit. These switches are progressively advanced at definite intervals, providing connection successively between each receiver and its corresponding transmitter. After a predetermined number of operations, a distinct synchronizing signal checks and, if necessary, adjusts the relative position of the switches at the transmitting and the receiving stations.

3.6 MEASUREMENT ERRORS

The complete statement of any measurement result has three elements: the *unit* in terms of which the result is stated; a *numeric* which states the magnitude of the result in terms of the chosen unit; and its *uncertainity,* the experimenter's estimate of the range within which the result may differ from the actual value of the quantity. Any physical measurement is uncertain to some extent, and errors are present in all phases of the measurement process, including the standards used to calibrate the system. Values assigned to local reference standards have uncertainties accumulated from the entire measurement chain extending back to the national reference standards that maintain a common measurement base. These national reference standards are themselves the experimental realization of the units defined in terms of the seven base units of the International System of Units (SI), and their assignments include an uncertainty estimate.

The Measurement Base. In most measurements, we are concerned only with their conformity within the technical community in which we work; our error chain stops at the national reference standards which maintain the *legal* units of the country, and our uncertainty estimates are based on these legal units. Rarely, when our concern is with the international measurement community or with basic science, must our uncertainty also include that of the maintained national unit.

Sources of Error. In addition to uncertainties in the calibration of a measurement system (which must be accepted as systematic errors in its operation), there are a number of error sources (some systematic and some random) in its operation. These operational error sources include noise, response time, design limitations, energy required by the system, signal transmission, system deterioration, and ambient influences.

Noise is any signal that does not convey measurement information. Disturbances generated within the system or coming from outside make up the background against which the desired signal must be read. Noise signals may be picked up by electrical or mechanical coupling between an external source and an element of the system, and may be amplified within the system. Under the most favorable circumstances, where noise has been minimized by filtering, by component selection, and by shielding and isolation of the system, there are still certain sources of noise present, resulting from the granular nature of matter and energy; the structure of phenomena is not infinitely fine-grained.

These fluctuations may be small compared with the total energy transfer involved in most measurements, yet they do give rise to a noise background that limits the ultimate sensitivity to which a measurement can be carried. Such sensitivity-limiting mechanisms include the brownian motion of a mechanical system, the Johnson noise in a resistance element, the Barkhausen effect in a magnetic element, and others.

The *response time* of a measuring system may contribute to measurement error. If the measured signal is not constant, lag in response results in an indication that depends on a sequence of values over a previous time interval.

Design limitations which contribute to measurement uncertainty include friction and resolution. Because a certain minimum force is needed to overcome friction and initiate motion, there results uncertainty in the rest position of an indicator. Resolution is the ability of the observer to distinguish between nearly equal quantities.

In an optical system, resolution is stated as the smallest angle at which points can be distinguished as separate. If the components of the optical train were perfect, resolution would be limited by the effective aperture of the system and the wavelength of the light used. If a scale is to be read to determine magnitudes, resolution is limited to the smallest fraction of a scale division that can be read with certainty. Most observers will attempt to estimate tenths of a division, but they generally have individual bias patterns that make a reading uncertain by 0.1 to 0.2 division.

Energy extracted from the measurand to operate the system alters the measurand to some extent, and if the available energy is small, this contributes to error in the result. Where energy is supplied from an auxiliary source, coupling or feedback may alter the measurement result. In the *transmission* of information from sensor to indicator, the signal may be distorted by selective attenuation or resonance in a communication channel, or it may suffer loss by leakage. Physical or chemical *deterioration* or other alterations of elements in the system can contribute to measurement error. Of the *ambient influences* affecting a measurement, temperature is the most pervasive. Other influences, not so universally important, include humidity, smoke and other air contaminants, barometric pressure, and the effect of gravity on an unbalanced system.

Classes of Errors. In estimating the uncertainty of a measurement result, two classes of error must be considered: systematic (which bias the result) and random (which produce scatter). *Systematic errors* are those which are repeated consistently with repetition of the measurement. Errors in the calibration of the system are systematic; uncertainty in the assigned value of a standard used in calibration must be accepted by the user as systematic. Changes of components through aging or deterioration produce systematic errors, as does failure to take into account energy extracted from a low-level source by the system. In attempting to search out and evaluate systematic errors, repetition of the measurement with definite, known changes in those parameters that are under the operator's control can be helpful, as is the use of different instrumentation or a different method. In some instances, it is possible to measure something similar to the measurand, which is independently and accurately known.

Random errors are accidental, fluctuating in an unpredictable manner. In any repetitive measurement, observations are influenced by many factors, the parameters that the observer cannot control and the residue of those he or she attempts to control. It is reasonable that combinations of these influences that add to produce large excursions are less frequent than those which partly compensate to produce small excursions, since each is equally likely to produce a positive or a negative departure. In effect, the results of a repetitive measurement process approximate a probability distribution, and it is convenient to treat the scatter of such a process as though it followed the laws of probability.

Evaluation Data. Assuming that the data from a repetitive measurement approximate a normal statistical distribution, we say that (excluding systematic errors) the mean of a group observations of a measurand is the *best approximation* of its actual value we can make from those data. Further, we estimate the imprecision of the result by certain statistical procedures. These procedures have validity only for random errors; systematic errors are not amenable to statistical treatment.

Standard Deviation. A measure of the dispersion of a set of observations is the root mean square of the deviations of individual observations from the mean of the set, $s = \sqrt{\Sigma d_m^2/(n - 1)}$, where n is the number of observations and d_m is the departure of an individual from the group mean. If the number is large, the standard deviation is $\sigma = \sqrt{\Sigma d_m^2/n}$. If the number of observations is small, a reasonable approximation of s can be calculated easily and quickly from the range r, the difference between the largest and smallest observation of the set $s \cong r/\sqrt{n}(3 \leq n \leq 12)$.

Probable Error. The probable error (pe) of an observation is that deviation from the mean for which the chances are equal that it will or will not be exceeded. If the number of observations is large, pe $= \pm 0.6745\sigma$. While this figure correctly expreses the range in which the chances are equally good that the actual value of the measurand will or will not be found (excluding systematics), it actually has no more significance as a precision index than the standard deviation from which it is derived. Thus, pe has fallen into disuse in current practice, although it was much used in the earlier literature as an index of precision. The pe of the mean of a set of observations is the amount by which the group mean can be expected to differ from the actual value of the measurand (excluding systematics) with a 50% probability. It may be calculated as $\pm\ 0.6745\sigma/\sqrt{n}$.

Confidence Intervals. Probable error is a special case of a broader concept. A *confidence interval* is the range of deviation from the mean within which a certain fraction of the observed values may be expected to lie, and the probability that the value of a randomly selected observation will lie within this range is called the *confidence level*.

BIBLIOGRAPHY

ANSIC12.1-2001: *Code for Electricity Metering.* New York: American National Standards Institute.

ASTM: *Manual on the Use of Thermocouples in Temperature Measurements.* Philadelphia: ASTM, 1981.

Beckwith, T. G.: *Mechanical Measurements.* Reading, Mass.: Addison-Wesley, 1993.

Benedict, Robert P.: *Fundamentals of Temperature and Pressure and Flow Measurements.* 3rd ed. New York: Wiley, 1984.

Berkeley Physics: *Electricity and Magnetism, vol. II*. New York: McGraw-Hill, 1985.

Doeblin, Ernest O.: *Measurement Systems and Design.* New York: McGraw-Hill, 1994.

Fowler, Richard J.: *Electricity Principles and Applications.* New York: McGraw-Hill, 1994.

Harris, Forest K.: *Electrical Measurements.* New York: Wiley, 1952.

IEC: International Electrotechnical Commission (IEC) Publication 751, Industrial Platinum Resistance Thermometer Sensors, Bureau Central de la Commission Electrotechnique Internationale. Geneva, Switzerland, 1983.

IEEE Transactions: *Instrumentation and Measurement* (*periodical*). New York: IEEE, 2005.

Keast, D. H.: *Measurements in Mechanical Dynamics.* New York: McGraw-Hill, 1967.

Keithley, Joseph, F.: *The Story of Electrical and Magnetic Measurements: From 500 B.C. to the 1940s.* New York: IEEE Press, 1998.

Mangum, B. W., and Furukawa, G. T.: *Guidelines for Realizing the International Temperature Scale of 1990,* NIST Technical Note 1265. Bethesda, Md.: NIST, 1990.

Thompson, Lawrence M.: *Electrical Measurements and Calibration: Fundamentals and Applications.* Research Triangle Park, N.C.: Instrument Society of America, 1994.

Tunbridge, Paul: *Lord Kelvin, His Influence on Electrical Measurements and Units.* London: P. Peregrinus, on behalf of the Institution of Electrical Engineers, 1992.

Webster, John G., ed.: *Electrical Measurement, Signal Processing, and Displays.* Boco Raton: CRC Press, 2003.

SECTION 4
PROPERTIES OF MATERIALS

Philip Mason Opsal
Wood Scientist, Wood Science LLC, Tucson, AZ

Grateful acknowledgement is also given to former contributors:

Donald J. Barta
Phelphs Dodge Company

T. W. Dakin
Westinghouse Research Laboratories

Charles A Harper
Technology Seminars, Inc.

Duane E. Lyon
Professor, Mississippi State University

Charles B. Rawlins
Alcoa Conductor Products

James Stubbins
Professor, University of Illinois

John Tanaka
Professor, University of Connecticut

CONTENTS

4.1 CONDUCTOR MATERIALS

4.1.1 General Properties

Conducting Materials. A *conductor of electricity* is any substance or material which will afford continuous passage to an electric current when subjected to a difference of electric potential. The greater the density of current for a given potential difference, the more efficient the conductor is said to be. Virtually, all substances in solid or liquid state possess the property of electric conductivity in some degree, but certain substances are relatively efficient conductors, while others are almost totally devoid of this property. The metals, for example, are the best conductors, while many other substances, such as metal oxides and salts, minerals, and fibrous materials, are relatively poor conductors, but their conductivity is beneficially affected by the absorption of moisture. Some of the less-efficient conducting materials such as carbon and certain metal alloys, as well as the efficient conductors such as copper and aluminum, have very useful applications in the electrical arts.

Certain other substances possess so little conductivity that they are classed as nonconductors, a better term being insulators or dielectrics. In general, all materials which are used commercially for conducting electricity for any purpose are classed as conductors.

Definition of Conductor. A *conductor* is a body so constructed from conducting material that it may be used as a carrier of electric current. In ordinary engineering usage, a conductor is a material of relatively high conductivity.

Types of Conductors. In general, a conductor consists of a solid wire or a multiplicity of wires stranded together, made of a conducting material and used either bare or insulated. Only bare conductors are considered in this subsection. Usually the conductor is made of copper or aluminum, but for applications requiring higher strength, such as overhead transmission lines, bronze, steel, and various composite constructions are used. For conductors having very low conductivity and used as resistor materials, a group of special alloys is available.

Definition of Circuit. An *electric circuit* is the path of an electric current, or more specifically, it is a conducting part or a system of parts through which an electric current is intended to flow. Electric circuits in general possess four fundamental electrical properties, consisting of resistance, inductance, capacitance, and leakage conductance. That portion of a circuit which is represented by its conductors will also possess these four properties, but only two of them are related to the properties of the conductor considered by itself. Capacitance and leakage conductance depend in part on the external dimensions of the conductors and their distances from one another and from other conducting bodies, and in part on the dielectric properties of the materials employed for insulating purposes. The inductance is a function of the magnetic field established by the current in a conductor, but this field as a whole is divisible into two parts, one being wholly external to the conductor and the other being wholly within the conductor; only the latter portion can be regarded as corresponding to the magnetic properties of the conductor material. The resistance is strictly a property of the conductor itself. Both the resistance and the internal inductance of conductors change in effective values when the current changes with great rapidity as in the case of high-frequency alternating currents; this is termed the *skin effect*.

In certain cases, conductors are subjected to various mechanical stresses. Consequently, their weight, tensile strength, and elastic properties require consideration in all applications of this character. Conductor materials as a class are affected by changes in temperature and by the conditions of mechanical stress to which they are subjected in service. They are also affected by the nature of the mechanical working and the heat treatment which they receive in the course of manufacture or fabrication into finished products.

4.1.2 Metal Properties

Specific Gravity and Density. Specific gravity is the ratio of mass of any material to that of the same volume of water at 4°C. Density is the unit weight of material expressed as pounds per cubic inch, grams per cubic centimeter, etc., at some reference temperature, usually 20°C. For all practical purposes, the numerical values of specific gravity and density are the same, expressed in g/cm^3.

Density and Weight of Copper. Pure copper, rolled, forged, or drawn and then annealed, has a density of 8.89 g/cm^3 at 20°C or 8.90 g/cm^3 at 0°C. Samples of high-conductivity copper usually will vary from 8.87 to 8.91 and occasionally from 8.83 to 8.94. Variations in density may be caused by microscopic flaws or seams or the presence of scale or some other defect; the presence of 0.03% oxygen will cause a reduction of about 0.01 in density. Hard-drawn copper has about 0.02% less density than annealed copper, on average, but for practical purposes the difference is negligible.

The international standard of density, 8.89 at 20°C, corresponds to a weight of 0.32117 lb/in^3 or 3.0270×10^{-6} lb/(cmil)(ft) or 15.982×10^{-3} lb/(cmil)(mile). Multiplying either of the last two figures by the square of the diameter of the wire in mils will produce the total weight of wire in pounds per foot or per mile, respectively.

Copper Alloys. Density and weight of copper alloys vary with the composition. For hard-drawn wire covered by ASTM Specification B105, the density of alloys 85 to 20 is 8.89 g/cm^3 (0.32117 lb/in^3) at 20°C; alloy 15 is 8.54 (0.30853); alloys 13 and 8.5 is 8.78 (0.31720).

Copper-Clad Steel. Density and weight of copper-clad steel wire is a mean between the density of copper and the density of steel, which can be calculated readily when the relative volumes or cross sections of copper and steel are known. For practical purposes, a value of 8.15 g/cm^3 (0.29444 lb/in^3) at 20°C is used.

Aluminum Wire. Density and weight of aluminum wire (commercially hard-drawn) is 2.705 g/cm^3 (0.0975 lb/in^3) at 20°C. The density of electrolytically refined aluminum (99.97% Al) and of hard-drawn wire of the same purity is 2.698 at 20°C. With less pure material there is an appreciable decrease in density on cold working. Annealed metal having a density of 2.702 will have a density of about 2.700 when in the hard-drawn or fully cold-worked conditions (see *NBS Circ*. 346, pp. 68 and 69).

Aluminum-Clad Wire. Density and weight of aluminum-clad wire is a mean between the density of aluminum and the density of steel, which can be calculated readily when the relative volumes or cross sections of aluminum and steel are known. For practical purposes, a value of 6.59 g/cm^3 (0.23808 lb/in^3) at 20°C is used.

Aluminum Alloys. Density and weight of aluminum alloys vary with type and composition. For hard-drawn aluminum alloy wire 5005-H19 and 6201-T81, a value of 2.703 g/cm^3 (0.09765 lb/in^3) at 20°C is used.

Pure Iron and Galvanized Steel Wire. Density and weight of pure iron is 7.90 g/cm^3 [2.690 $\times$ 10^{-6} lb/(cmil)(ft)] at 20°C. Density and weight of galvanized steel wire (EBB, BB, HTL-85, HTL-135, and HTL-195) with Class A weight of zinc coating are 7.83 g/cm^3 (0.283 lb/in^3) at 20°C, with Class B are 7.80 g/cm^3 (0.282 lb/in^3), and with Class C are 7.78 g/cm^3 (0.281 lb/in^3).

Percent Conductivity. It is very common to rate the conductivity of a conductor in terms of its percentage ratio to the conductivity of chemically pure metal of the same kind as the conductor is primarily constituted or in ratio to the conductivity of the international copper standard. Both forms of the conductivity ratio are useful for various purposes. This ratio also can be expressed in two different terms, one where the conductor cross sections are equal and therefore termed the *volume-conductivity ratio* and the other where the conductor masses are equal and therefore termed the *mass-conductivity ratio*.

International Annealed Copper Standard. The International Annealed Copper Standard (IACS) is the internationally accepted value for the resistivity of annealed copper of 100% conductivity. This standard is expressed in terms of mass resistivity as 0.5328 $\Omega \cdot$ g/m^2, or the resistance of a uniform round wire 1 m long weighing 1 g at the standard temperature of 20°C. Equivalent expressions of the annealed copper standard in various units of mass resistivity and volume resistivity are as follows:

0.15328	$\Omega \cdot$ g/m^2
875.20	$\Omega \cdot$ lb/mi^2
1.7241	$\mu\Omega \cdot$ cm
0.67879	$\mu\Omega \cdot$ in at 20°C
10.371	$\Omega \cdot$ cmil/ft
0.017241	$\Omega \cdot$ mm^2/m

The preceding values are the equivalent of $^1/_{58}$ $\Omega \cdot$ mm^2/m, so the volume conductivity can be expressed as 58 S $\cdot$ mm^2/m at 20°C.

Conductivity of Conductor Materials. Conductivity of conductor materials varies with chemical composition and processing.

Electrical Resistivity. *Electrical resistivity* is a measure of the resistance of a unit quantity of a given material. It may be expressed in terms of either mass or volume; mathematically,

Mass resistivity:
$$\delta = \frac{Rm}{l^2} \tag{4-1}$$

Volume resistivity:
$$\rho = \frac{RA}{l} \tag{4-2}$$

where R is resistance, m is mass, A is cross-sectional area, and l is length.

Electrical resistivity of conductor materials varies with chemical composition and processing.

Effects of Temperature Changes. Within the temperature ranges of ordinary service there is no appreciable change in the properties of conductor materials, except in electrical resistance and physical dimensions. The change in resistance with change in temperature is sufficient to require consideration in many engineering calculations. The change in physical dimensions with change in temperature is also important in certain cases, such as in overhead spans and in large units of apparatus or equipment.

Temperature Coefficient of Resistance. Over moderate ranges of temperature, such as 100°C, the change of resistance is usually proportional to the change of temperature. Resistivity is always expressed at a standard temperature, usually 20°C (68°F). In general, if R_{t_1} is the resistance at a temperature t_1

and α_{t_1} is the temperature coefficient at that temperature, the resistance at some other temperature t_2 is expressed by the formula

$$R_{t_2} = R_{t_1}[1 + \alpha_{t_1}(t_2 - t_1)] \tag{4-3}$$

Over wide ranges of temperature, the linear relationship of this formula is usually not applicable, and the formula then becomes a series involving higher powers of t, which is unwieldy for ordinary use.

When the temperature of reference t_1 is changed to some other value, the coefficient also changes. Upon assuming the general linear relationship between resistance and temperature previously mentioned, the new coefficient at any temperature t within the linear range is expressed

$$\alpha_t = \frac{1}{(1/\alpha_{t_1}) + (t - t_1)} \tag{4-4}$$

The reciprocal of α is termed the *inferred absolute zero of temperature*. Equation (4-3) takes no account of the change in dimensions with change in temperature and therefore applies to the case of conductors of constant mass, usually met in engineering work.

The *coefficient for copper of less than standard (or 100%) conductivity* is proportional to the actual conductivity, expressed as a decimal percentage. Thus, if n is the percentage conductivity (95% = 0.95), the temperature coefficient will be $\alpha'_t = n\alpha_t$, where α_t is the coefficient of the annealed copper standard.

The coefficients are computed from the formula

$$\alpha_t = \frac{1}{[1/n(0.00393)] + (t_1 - 20)} \tag{4-5}$$

Copper Alloys and Copper-Clad Steel Wire. Temperature-resistance coefficients for copper alloys usually can be approximated by multiplying the corresponding coefficient for copper (100% IACS) by the alloy conductivity expressed as a decimal. For some complex alloys, however, this relation does not hold even approximately, and suitable values should be obtained from the supplier. The temperature-resistance coefficient for copper-clad steel wire is 0.00378/°C at 20°C.

Aluminum-Alloy Wires and Aluminum-Clad Wire. Temperature-resistance coefficients for aluminum-alloy wires are for 5005 H19, 0.00353/°C, and for 6201-T81, 0.00347/°C at 20°C. Temperature-resistance coefficient for aluminum-clad wire is 0.0036/°C at 20°C.

Typical Composite Conductors. Temperature-resistance coefficients for typical composite conductors are as follows:

Type	Approximate temperature coefficient per °C at 20°C
Copper–copper-clad steel	0.00381
ACSR (aluminum-steel)	0.00403
Aluminum–aluminum alloy	0.00394
Aluminum–aluminum-clad steel	0.00396

Reduction of Observations to Standard Temperature. A table of convenient corrections and factors for reducing resistivity and resistance to standard temperature, 20°C, will be found in Copper Wire Tables, *NBS Handbook* 100.

Resistivity-Temperature Constant. The *change of resistivity per degree* may be readily calculated, taking account of the expansion of the metal with rise of temperature. The proportional relation between temperature coefficient and conductivity may be put in the following convenient form for reducing *resistivity* from one temperature to another. *The change of resistivity of copper per degree Celsius is a constant, independent of the temperature of reference and of the sample of copper. This "resistivity-temperature constant" may be taken, for general purposes, as* 0.00060 Ω *(meter, gram), or* 0.0068 μΩ · *cm.*

Details of the calculation of the resistivity-temperature constant will be found in Copper Wire Tables, *NBS Handbook* 100; also see this reference for expressions for the temperature coefficients of resistivity and their derivation.

Temperature Coefficient of Expansion. *Temperature coefficient of expansion* (linear) of pure metals over a range of several hundred degrees is not a linear function of the temperature but is well expressed by a quadratic equation

$$\frac{L_{t_2}}{L_{t_1}} = 1 + [\alpha(t_2 - t_1) + \beta(t_2 - t_1)^2] \tag{4-6}$$

Over the temperature ranges for ordinary engineering work (usually 0 to 100°C), the coefficient can be taken as a constant (assumed linear relationship) and a simplified formula employed

$$L_{t_2} = L_{t_1}[1 + \alpha_{t_1}(t_2 - t_1)] \tag{4-7}$$

Changes in linear dimensions, superficial area, and volume take place in most materials with changes in temperature. In the case of linear conductors, only the change in length is ordinarily important.

The coefficient for changes in superficial area is approximately twice the coefficient of linear expansion for relatively small changes in temperature. Similarly, the volume coefficient is 3 times the linear coefficient, with similar limitations.

Specific Heat. Specific heat of electrolytic tough pitch copper is 0.092 cal/(g)(°C) at 20°C (see *NBS Circ.* 73). Specific heat of aluminum is 0.226 cal/(g)(°C) at room temperature (see *NBS Circ.* C447, Mechanical Properties of Metals and Alloys). Specific heat of iron (wrought) or very soft steel from 0 to 100°C is 0.114 cal/(g)(°C); the true specific heat of iron at 0°C is 0.1075 cal/(g)(°C) (see *International Critical Tables,* vol. II, p. 518; also ASM, *Metals Handbook*).

Thermal Conductivity of Electrolytic Tough Pitch Copper. Thermal conductivity of electrolytic tough pitch copper at 20°C is 0.934 cal/(cm^2)(cm)(s)(°C), adjusted to correspond to an electrical conductivity of 101% (see *NBS Circ.* 73).

Thermal-Electrical Conductivity Relation of Copper. The Wiedemann-Franz-Lorenz law, which states that the ratio of the thermal and electrical conductivities at a given temperature is independent of the nature of the conductor, holds closely for copper. The ratio $K/\lambda T$ (where K = thermal conductivity, λ = electrical conductivity, T = absolute temperature) for copper is 5.45 at 20°C.

Thermal Conductivity.

Copper Alloys.

	Thermal conductivity (volumetric) at 20°C	
ASTM alloy (Spec. B105)	Btu per sq ft per ft per h per °F	Cal per sq cm per cm per sec per °C
8.5	31	0.13
15	50	0.21
30	84	0.35
55	135	0.56
80	199	0.82
85	208	0.86

Aluminum. The determination made by the Bureau of Standards at 50°C for aluminum of 99.66% purity is 0.52 cal/(cm^2)(cm)(s)(°C) (*Circ*. 346; also see *Smithsonian Physical Tables and International Critical Tables*).

Iron. Thermal conductivity of iron (mean) from 0 to 100°C is 0.143 cal/(cm^2)(cm)(s)(°C); with increase of carbon and manganese content, it tends to decrease and may reach a figure of approximately

0.095 with about 1% carbon, or only about half of that figure if the steel is hardened by water quenching (see *International Critical Tables,* vol. II, p. 518).

Copper. Copper is a highly malleable and ductile metal of reddish color. It can be cast, forged, rolled, drawn, and machined. Mechanical working hardens it, but annealing will restore it to the soft state. The density varies slightly with the physical state, 8.9 being an average value. It melts at 1083°C (1981°F) and in the molten state has a sea-green color. When heated to a very high temperature, it vaporizes and burns with a characteristic green flame. Copper readily alloys with many other metals. In ordinary atmospheres it is not subject to appreciable corrosion. Its electrical conductivity is very sensitive to the presence of slight impurities in the metal.

Copper, when exposed to ordinary atmospheres, becomes oxidized, turning to a black color, but the oxide coating is protective, and the oxidizing process is not progressive. When exposed to moist air containing carbon dioxide, it becomes coated with green basic carbonate, which is also protective. At temperatures above 180°C it oxidizes in dry air. In the presence of ammonia it is readily oxidized in air, and it is also affected by sulfur dioxide. Copper is not readily attacked at high temperatures below the melting point by hydrogen, nitrogen, carbon monoxide, carbon dioxide, or steam. Molten copper readily absorbs oxygen, hydrogen, carbon monoxide, and sulfur dioxide, but on cooling, the occluded gases are liberated to a great extent, tending to produce blowholes or porous castings. Copper in the presence of air does not dissolve in dilute hydrochloric or sulfuric acid but is readily attacked by dilute nitric acid. It is also corroded slowly by saline solutions and sea water.

Commercial grades of copper in the United States are electrolytic, oxygen-free, Lake, fire-refined, and casting. *Electrolytic copper* is that which has been electrolytically refined from blister, converter, black, or Lake copper. *Oxygen-free copper* is produced by special manufacturing processes which prevent the absorption of oxygen during the melting and casting operations or by removing the oxygen by reducing agents. It is used for conductors subjected to reducing gases at elevated temperature, where reaction with the included oxygen would lead to the development of cracks in the metal. *Lake copper* is electrolytically or fire-refined from Lake Superior native copper ores and is of two grades, low resistance and high resistance. *Fire-refined copper* is a lower-purity grade intended for alloying or for fabrication into products for mechanical purposes; it is not intended for electrical purposes. *Casting copper* is the grade of lowest purity and may consist of furnace-refined copper, rejected metal not up to grade, or melted scrap; it is exclusively a foundry copper.

Hardening and Heat-Treatment of Copper. There are but two well-recognized methods for hardening copper, one is by mechanically working it, and the other is by the addition of an alloying element. The properties of copper are not affected by a rapid cooling after annealing or rolling, as are those of steel and certain copper alloys.

Annealing of Copper. Cold-worked copper is softened by annealing, with decrease of tensile strength and increase of ductility. In the case of pure copper hardened by cold reduction of area to one-third of its initial area, this softening takes place with maximum rapidity between 200 and 325°C. However, this temperature range is affected in general by the extent of previous cold reduction and the presence of impurities. The greater the previous cold reduction, the lower is the range of softening temperatures. The effect of iron, nickel, cobalt, silver, cadmium, tin, antimony, and tellurium is to lower the conductivity and raise the annealing range of pure copper in varying degrees.

Commercial grade	ASTM Designation	Copper content, minimum %
Electrolytic	B5	99.900
Oxygen-free electrolytic	B170	99.95
Lake, low resistance	B4	99.900
Lake, high resistance	B4	99.900
Fire-refined	B216	99.88
Casting	B119	98

Alloying of Copper. Elements that are soluble in moderate amounts in a solid solution of copper, such as manganese, nickel, zinc, tin, and aluminum, generally harden it and diminish its ductility but improve its rolling and working properties. Elements that are but slightly soluble, such as bismuth and lead, do not harden it but diminish both the ductility and the toughness and impair its hot-working properties. Small additions (up to 1.5%) of manganese, phosphorus, or tin increase the tensile strength and hardness of cold-rolled copper.

Brass is usually a binary alloy of copper and zinc, but brasses are seldom employed as electrical conductors, since they have relatively low conductivity through comparatively high tensile strength. In general, brass is not suitable for use when exposed to the weather, owing to the difficulty from stress-corrosion cracking; the higher the zinc content, the more pronounced this becomes.

Bronze in its simplest form is a binary alloy of copper and tin in which the latter element is the hardening and strengthening agent. This material is rather old in the arts and has been used to some extent for electrical conductors for past many years, especially abroad. Modern bronzes are frequently ternary alloys, containing as the third constituent such elements as phosphorus, silicon, manganese, zinc, aluminum, or cadmium; in such cases, the third element is usually given in the name of the alloy, as in phosphor bronze or silicon bronze. Certain bronzes are quaternary alloys, or contain two other elements in addition to copper and tin.

In bronzes for use as electrical conductors, the content of tin and other metals is usually less than in bronzes for structural or mechanical applications, where physical properties and resistance to corrosion are the governing considerations. High resistance to atmospheric corrosion is always an important consideration in selecting bronze conductors for overhead service.

Commercial Grades of Bronze. Various bronzes have been developed for use as conductors, and these are now covered by ASTM Specification B105. They all have been designed to provide conductors having high resistance to corrosion and tensile strengths greater than hard-drawn copper conductors. The standard specification covers 10 grades of bronze, designated by numbers according to their conductivities.

Copper-Beryllium Alloy. *Copper-beryllium alloy* containing 0.4% of beryllium may have an electrical conductivity of 48% and a tensile strength (in 0.128-in wire) of 86,000 lb/in^2. A content of 0.9% of beryllium may give a conductivity of 28% and a tensile strength of 122,000 lb/in^2. The effect of this element in strengthening copper is about 10 times as great as that of tin.

Copper-Clad Steel Wire. *Copper-clad steel wire* has been manufactured by a number of different methods. The general object sought in the manufacture of such wires is the combination of the high conductivity of copper with the high strength and toughness of iron or steel. The principal manufacturing processes now in commercial use are (*a*) coating a steel billet with a special flux, placing it in a vertical mold closed at the bottom, heating the billet and mold to yellow heat, and then casting molten copper around the billet, after which it is hot-rolled to rods and cold-drawn to wire, and (*b*) electroplating a dense coating of copper on a steel rod and then cold drawing to wire.

Aluminum. *Aluminum* is a ductile metal, silver-white in color, which can be readily worked by rolling, drawing, spinning, extruding, and forging. Its specific gravity is 2.703. Pure aluminum melts at 660°C (1220°F). Aluminum has relatively high thermal and electrical conductivities. The metal is always covered with a thin, invisible film of oxide which is impermeable and protective in character. Aluminum, therefore, shows stability and long life under ordinary atmospheric exposure.

Exposure to atmospheres high in hydrogen sulfide or sulfur dioxide does not cause severe attack of aluminum at ordinary temperatures, and for this reason, aluminum or its alloys can be used in atmospheres which would be rapidly corrosive to many other metals.

Aluminum parts should, as a rule, not be exposed to salt solutions while in electrical contact with copper, brass, nickel, tin, or steel parts, since galvanic attack of the aluminum is likely to occur. Contact with cadmium in such solutions results in no appreciable acceleration in attack on the aluminum, while contact with zinc (or zinc-coated steel as long as the coating is intact) is generally beneficial, since the zinc is attacked selectively and it cathodically protects adjacent areas of the aluminum.

Most organic acids and their water solutions have little or no effect on aluminum at room temperature, although oxalic acid is an exception and is corrosive. Concentrated nitric acid (about 80% by weight) and fuming sulfuric acid can be handled in aluminum containers. However, more dilute solutions of these acids are more active. All but the most dilute (less than 0.1%) solutions of hydrochloric and hydrofluoric acids have a rapid etching action on aluminum.

Solutions of the strong alkalies, potassium, or sodium hydroxides dissolve aluminum rapidly. However, ammonium hydroxide and many of the strong organic bases have little action on aluminum and are successfully used in contact with it (see *NBS Circ.* 346).

Aluminum in the presence of water and limited air or oxygen rapidly converts into aluminum hydroxide, a whitish powder.

Commercial grades of aluminum in the United States are designated by their purity, such as 99.99, 99.6, 99.2, 99.0%. Electrical conductor alloy aluminum 1350, having a purity of approximately 99.5% and a minimum conductivity of 61.0% IACS, is used for conductor purposes. Specified physical properties are obtained by closely controlling the kind and amount of certain impurities.

Annealing of Aluminum. Cold-worked aluminum is softened by annealing, with decrease of tensile strength and increase of ductility. The annealing temperature range is affected in general by the extent of previous cold reduction and the presence of impurities. The greater the previous cold reduction, the lower is the range of softening temperatures.

Alloying of Aluminum. Aluminum can be alloyed with a variety of other elements, with a consequent increase in strength and hardness. With certain alloys, the strength can be further increased by suitable heat treatment. The alloying elements most generally used are copper, silicon, manganese, magnesium, chromium, and zinc. Some of the aluminum alloys, particularly those containing one or more of the following elements—copper, magnesium, silicon, and zinc—in various combinations, are susceptible to heat treatment.

Pure aluminum, even in the hard-worked condition, is a relatively weak metal for construction purposes. Strengthening for castings is obtained by alloying elements. The alloys most suitable for cold rolling seldom contain less than 90% to 95% aluminum. By alloying, working, and heat treatment, it is possible to produce tensile strengths ranging from 8500 lb/in^2 for pure annealed aluminum up to 82,000 lb/in^2 for special wrought heat-treated alloy, with densities ranging from 2.65 to 3.00.

Electrical conductor alloys of aluminum are principally alloys 5005 and 6201 covered by ASTM Specifications B396 and B398.

Aluminum-clad steel wires have a relatively heavy layer of aluminum surrounding and bonded to the high-strength steel core. The aluminum layer can be formed by compacting and sintering a layer of aluminum powder over a steel rod, by electroplating a dense coating of aluminum on a steel rod, or by extruding a coating of aluminum on a steel rod and then cold drawing to wire.

Silicon. *Silicon* is a light metal having a specific gravity of approximately 2.34. There is lack of accurate data on the pure metal because its mechanical brittleness bars it from most industrial uses. However, it is very resistant to atmospheric corrosion and to attack by many chemical reagents. Silicon is of fundamental importance in the steel industry, but for this purpose it is obtained in the form of ferrosilicon, which is a coarse granulated or broken product. It is very useful as an alloying element in steel for electrical sheets and substantially increases the electrical resistivity, and thereby reduces the core losses. Silicon is peculiar among metals in the respect that its temperature coefficient of resistance may change sign in some temperature ranges, the exact behavior varying with the impurities.

Beryllium. *Beryllium* is a light metal having a specific gravity of approximately 1.84, or nearly the same as magnesium. It is normally hard and brittle and difficult to fabricate. Copper is materially strengthened by the addition of small amounts of beryllium, without very serious loss of electrical conductivity. The principal use for this metal appears to be as an alloying element with other metals such as aluminum and copper. Beryllium is also toxic. Reference should be made to Material Safety Data Sheets for precautions in handling.

Sodium. *Sodium* is a soft, bright, silvery metal obtained commercially by the electrolysis of absolutely dry fused sodium chloride. It is the most abundant of the alkali group of metals, is extremely reactive, and is never found free in nature. It oxidizes readily and rapidly in air. In the presence of water (it is so light that it floats) it may ignite spontaneously, decomposing the water with evolution of hydrogen and formation of sodium hydroxide. This can be explosive. Sodium should be handled with respect, since it can be dangerous when handled improperly. It melts at 97.8°C, below the boiling point of water and in the same range as many fuse metal alloys. Sodium is approximately one-tenth as heavy as copper and has roughly three-eighths the conductivity; hence 1 lb of sodium is about equal electrically to 3½ lb of copper.

4.1.3 Conductor Properties

Definitions of Electrical Conductors

Wire. A rod or filament of drawn or rolled metal whose length is great in comparison with the major axis of its cross section. The definition restricts the term to what would ordinarily be understood by the term *solid wire*. In the definition, the word *slender* is used in the sense that the length is great in comparison with the diameter. If a wire is covered with insulation, it is properly called an *insulated wire,* while primarily the term *wire* refers to the metal; nevertheless, when the context shows that the wire is insulated, the term *wire* will be understood to include the insulation.

Conductor. A wire or combination of wires not insulated from one another, suitable for carrying an electric current. The term *conductor* is not to include a combination of conductors insulated from one another, which would be suitable for carrying several different electric currents. Rolled conductors (such as bus bars) are, of course, conductors but are not considered under the terminology here given.

Stranded Conductor. A conductor composed of a group of wires, usually twisted, or any combination of groups of wires. The wires in a stranded conductor are usually twisted or braided together.

Cable. A stranded conductor (single-conductor cable) or a combination of conductors insulated from one another (multiple-conductor cable). The component conductors of the second kind of cable may be either solid or stranded, and this kind of cable may or may not have a common insulating covering. The first kind of cable is a single conductor, while the second kind is a group of several conductors. The term *cable* is applied by some manufacturers to a solid wire heavily insulated and lead covered; this usage arises from the manner of the insulation, but such a conductor is not included under this definition of *cable*. The term *cable* is a general one, and in practice, it is usually applied only to the larger sizes. A small cable is called a *stranded wire* or a *cord,* both of which are defined below. Cables may be bare or insulated, and the latter may be armored with lead or with steel wires or bands.

Strand. One of the wires of any stranded conductor.

Stranded Wire. A group of small wires used as a single wire. A *wire* has been defined as a slender rod or filament of drawn metal. If such a filament is subdivided into several smaller filaments or strands and is used as a single wire, it is called *stranded wire*. There is no sharp dividing line of size between a stranded wire and a cable. If used as a wire, for example, in winding inductance coils or magnets, it is called a *stranded wire* and not a cable. If it is substantially insulated, it is called a *cord,* defined below.

Cord. A small cable, very flexible and substantially insulated to withstand wear. There is no sharp dividing line in respect to size between a cord and a cable, and likewise no sharp dividing line in respect to the character of insulation between a cord and a stranded wire. Usually the insulation of a cord contains rubber.

Concentric Strand. A strand composed of a central core surrounded by one or more layers of helically laid wires or groups of wires.

Concentric-Lay Conductor. Conductor constructed with a central core surrounded by one or more layers of helically laid wires.

Rope-Lay Conductor. Conductor constructed of a bunch-stranded or a concentric-stranded member or members, as a central core, around which are laid one or more helical layers of such members.

N-Conductor Cable. A combination of *N* conductors insulated from one another. It is not intended that the name as given here actually be used. One would instead speak of a "3-conductor cable," a "12-conductor cable," etc. In referring to the general case, one may speak of a "multiple-conductor cable."

N-Conductor Concentric Cable. A cable composed of an insulated central conducting core with *N*-1 tubular-stranded conductors laid over it concentrically and separated by layers of insulation. This kind of cable usually has only two or three conductors. Such cables are used in carrying alternating currents. The remark on the expression "*N* conductor" given for the preceding definition applies here also. (Additional definitions can be found in ASTM B354.)

Wire Sizes. Wire sizes have been for many years indicated in commercial practice almost entirely by gage numbers, especially in America and England. This practice is accompanied by some confusion because numerous gages are in common use. The most commonly used gage for electrical wires, in America, is the *American wire gage*. The most commonly used gage for steel wires is the *Birmingham wire gage*.

There is no legal standard wire gage in this country, although a gage for sheets was adopted by Congress in 1893. In England, there is a legal standard known as the *Standard wire gage*. In Germany, France, Austria, Italy, and other continental countries, practically no wire gage is used, but wire sizes are specified directly in millimeters. This system is sometimes called the *millimeter wire gage*. The wire sizes used in France, however, are based to some extent on the old Paris gage (*jauge de Paris de 1857*) (for a history of wire gages, see *NBS Handbook* 100, Copper Wire Tables; also see *Circ.* 67, Wire Gages, 1918).

There is a tendency to *abandon gage numbers* entirely and specify wire sizes by the *diameter in mils* (thousandths of an inch). This practice holds particularly in writing specifications and has the great advantages of being both simple and explicit. A number of wire manufacturers also encourage this practice, and it was definitely adopted by the U.S. Navy Department in 1911.

Mil is a term universally employed in this country to measure wire diameters and is a unit of length equal to one-thousandth of an inch. *Circular mil* is a term universally used to define cross-sectional areas, being a unit of area equal to the area of a circle 1 mil in diameter. Such a circle, however, has an area of 0.7854 (or $\pi/4$) mil^2. Thus a wire 10 mils in diameter has a cross-sectional area of 100 cmils or 78.54 mils^2. Hence, a cmil equals 0.7854 mil^2.

American wire gage, also known as the *Brown & Sharpe gage,* was devised in 1857 by J. R. Brown. It is usually abbreviated AWG. This gage has the property, in common with a number of other gages, that its sizes represent approximately the successive steps in the process of wire drawing. Also, like many other gages, its numbers are retrogressive, a larger number denoting a smaller wire, corresponding to the operations of drawing. These gage numbers are not arbitrarily chosen, as in many gages, but follow the mathematical law upon which the gage is founded.

Basis of the AWG is a simple mathematical law. The gage is formed by the specification of two diameters and the law that a given number of intermediate diameters are formed by geometric progression. Thus, the diameter of No. 0000 is defined as 0.4600 in and of No. 36 as 0.0050 in. There are 38 sizes between these two; hence the ratio of any diameter to the diameter of the next greater number is given by this expression

$$\sqrt[39]{\frac{0.4600}{0.0050}} = \sqrt[39]{92} = 1.122\,932\,2 \tag{4-8}$$

The square of this ratio = 1.2610. The sixth power of the ratio, that is, the ratio of any diameter to the diameter of the sixth greater number, = 2.0050. The fact that this ratio is so nearly 2 is the basis of numerous useful relations or shortcuts in wire computations.

There are a number of approximate rules applicable to the AWG which are useful to remember:

1. An increase of three gage numbers (e.g., from No. 10 to 7) doubles the area and weight and consequently halves the dc resistance.
2. An increase of six gage numbers (e.g., from No. 10 to 4) doubles the diameter.
3. An increase of 10 gage numbers (e.g., from No. 10 to 1/0) multiplies the area and weight by 10 and divides the resistance by 10.
4. A No. 10 wire has a diameter of about 0.10 in, an area of about 10,000 cmils, and (for standard annealed copper at 20°C) a resistance of approximately 1.0 Ω/1000 ft.
5. The weight of No. 2 copper wire is very close to 200 lb/1000 ft (90 kg/304.8 m).

Steel wire gage, also known originally as the *Washburn & Moen gage* and later as the *American Steel & Wire Co.'s gage,* was established by Ichabod Washburn in 1830. This gage, with a number of its sizes rounded off to thousandths of an inch, is also known as the *Roebling gage*. It is used exclusively for steel wire and is frequently employed in wire mills.

Birmingham wire gage, also known as *Stubs' wire gage* and *Stubs' iron wire gage,* is said to have been established early in the eighteenth century in England, where it was long in use. This gage was used to designate the Stubs soft-wire sizes and should not be confused with Stubs' steel-wire gage. The numbers of the Birmingham gage were based on the reductions of size made in practice by drawing wire from rolled rod. Thus, a wire rod was called "No. 0," "first drawing No. 1," and so on. The gradations of size in this gage are not regular, as will appear from its graph. This gage is generally in commercial use in the United States for iron and steel wires.

Standard wire gage, which more properly should be designated *(British) Standard wire gage,* is the legal standard of Great Britain for all wires adopted in 1883. It is also known as the *New British Standard gage,* the *English legal standard gage,* and the *Imperial wire gage*. It was constructed by so modifying the Birmingham gage that the differences between consecutive sizes become more regular. This gage is largely used in England but never has been used extensively in America.

Old English wire gage, also known as the *London wire gage,* differs very little from the Birmingham gage. Formerly it was used to some extent for brass and copper wires but is now nearly obsolete.

Millimeter wire gage, also known as the *metric wire gage*, is based on giving progressive numbers to the progressive sizes, calling 0.1 mm diameter "No. 1," 0.2 mm "No. 2," etc.

Conductor-Size Designation. America uses, for sizes up to 4/0, mil, decimals of an inch, or AWG numbers for solid conductors and AWG numbers or circular mils for stranded conductors; for sizes larger than 4/0, circular mils are used throughout. Other countries ordinarily use square millimeter area.

Conductor-size conversion can be accomplished from the following relation:

$$\text{cmils} = \text{in}^2 \times 1{,}273{,}200 = \text{mm}^2 \times 1973.5 \tag{4-9}$$

Measurement of wire diameters may be accomplished in many ways but most commonly by means of a micrometer caliper. Stranded cables are usually measured by means of a circumference tape calibrated directly in diameter readings.

Stranded Conductors. Stranded conductors are used generally because of their increased flexibility and consequent ease in handling. The greater the number of wires in any given cross section, the greater will be the flexibility of the finished conductor. Most conductors above 4/0 AWG in size are stranded. Generally, in a given concentric-lay stranded conductor, all wires are of the same size and the same material, although special conductors are available embodying wires of different sizes and materials. The former will be found in some insulated cables and the latter in overhead stranded conductors combining high-conductivity and high-strength wires.

The flexibility of any given size of strand obviously increases as the total number of wires increases. It is a common practice to increase the total number of wires as the strand diameter increases in order to provide reasonable flexibility in handling. So-called flexible concentric strands for use in insulated cables have about one to two more layers of wires than the standard type of strand for ordinary use.

Number of Wires in Standard Conductors. Each successive layer in a concentrically stranded conductor contains six more wires than the preceding one. The total number of wires in a conductor is

For 1-wire core constructions (1, 7, 19, etc.),

$$N = 3n(n + 1) + 1 \tag{4-10}$$

For 3-wire core constructions (3, 12, etc.),

$$N = 3n(n + 2) + 3 \tag{4-11}$$

where n is number of layers over core, which is not counted as a layer.

Wire size in stranded conductors is

$$d = \sqrt{\frac{A}{N}} \tag{4-12}$$

where A is total conductor area in circular mils, and N is total number of wires.

Copper cables are manufactured usually to certain cross-sectional sizes specified in total circular mils or by gage numbers in AWG. This necessarily requires individual wires drawn to certain prescribed diameters, which are different as a rule from normal sizes in AWG (see Table 4-10).

Diameter of stranded conductors (circumscribing circle) is

$$D = d(2n + k) \tag{4-13}$$

where d is diameter of individual wire, n is number of layers over core, which is not counted as a layer, k is 1 for constructions having 1-wire core (1, 7, 19, etc.), and 2.155 for constructions having 3-wire core (3, 12, etc.).

For standard concentric-lay stranded conductors, the following rule gives a simple method of determining the outside diameter of a stranded conductor from the known diameter of a solid wire of the same cross-sectional area: *To obtain the diameter of concentric-lay stranded conductor, multiply the diameter of the solid wire of the same cross-sectional area by the appropriate factor as follows*:

Number of wires	Factor	Number of wires	Factor
3	1.244	91	1.153
7	1.134	127	1.154
12	1.199	169	1.154
19	1.147	217	1.154
37	1.151	271	1.154
61	1.152		

Area of stranded conductors is

$$A = Nd^2 \text{ cmils} = \tfrac{1}{4}\,\pi N d^2 \times 10^{-6} \text{ in}^2 \tag{4-14}$$

where N is total number of wires, and d is individual wire diameter in mils.

Effects of Stranding. All wires in a stranded conductor except the core wire form continuous helices of slightly greater length than the axis or core. This causes a slight increase in weight and electrical resistance and slight decrease in tensile strength and sometimes affects the internal inductance, as compared theoretically with a conductor of equal dimensions but composed of straight wires parallel with the axis.

Lay, or Pitch. The axial length of one complete turn, or helix, of a wire in a stranded conductor is sometimes termed the *lay,* or *pitch*. This is often expressed as the *pitch ratio,* which is the ratio of the length of the helix to its *pitch diameter* (diameter of the helix at the centerline of any individual wire or strand equals the outside diameter of the helix minus the thickness of one wire or strand). If there are several layers, the pitch expressed as an axial length may increase with each additional layer, but when expressed as the ratio of axial length to pitch diameter of helix, it is usually the same for all layers, or nearly so. In commercial practice, the pitch is commonly expressed as the ratio of axial length to outside diameter of helix, but this is an arbitrary designation made for convenience of usage. The *pitch angle* is shown in Fig. 4-1, where *ac* represents the axis of the stranded conductor and l is the axial length of one complete turn or helix, *ab* is the length of any individual wire $l + \Delta l$ in one complete turn, and *bc* is equal to the circumference of a circle corresponding to the pitch diameter d of the helix. The angle *bac,* or θ, is the pitch angle, and the pitch ratio is expressed by $p = l/d$. There is no standard pitch ratio used by manufacturers generally, since it has been found desirable to vary this depending on the type of service for which the conductor is intended. Applicable lay lengths generally are included in industry specifications covering the various stranded conductors. For bare overhead conductors, a representative commercial value for pitch length is 13.5 times the outside diameter of each layer of strands.

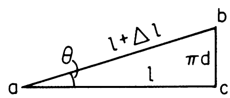


FIGURE 4-1 Pitch angle in concentric-lay cable.

Direction of Lay. The direction of lay is the lateral direction in which the individual wires of a cable run over the top of the cable as they recede from an observer looking along the axis. *Right-hand lay* recedes from the observer in clockwise rotation or like a right-hand screw thread; *left-hand lay* is the opposite. The outer layer of a cable is ordinarily applied with a right-hand lay for bare overhead conductors and left-hand lay for insulated conductors, although the opposite lay can be used if desired.

Increase in Weight Due to Stranding. Referring to Fig. 4-1, the increase in weight of the spiral members in a cable is proportional to the increase in length

$$\frac{l + \Delta l}{l} = \sec\theta = \sqrt{1 + \tan^2\theta}$$

$$= \sqrt{1 + \frac{\pi^2}{p^2}} = 1 + \frac{1}{2}\frac{\pi^2}{p^2} - \frac{1}{8}\left(\frac{\pi^2}{p^2}\right)^2 + \cdots \qquad (4\text{-}15)$$

As a first approximation this ratio equals $1 + 0.5(\pi^2/p^2)$, and a pitch of 15.7 produces a ratio of 1.02. This correction factor should be computed separately for each layer if the pitch p varies from layer to layer.

Increase in Resistance Due to Stranding. If it were true that no current flows from wire to wire through their lineal contacts, the proportional increase in the total resistance would be the same as the proportional increase in total weight. If all the wires were in perfect and complete contact with each other, the total resistance would decrease in the same proportion that the total weight increases, owing to the slightly increased normal cross section of the cable as a whole. The contact resistances are normally sufficient to make the actual increase in total resistance nearly as much, proportionately, as the increase in total weight, and for practical purposes they are usually assumed to be the same.

Decrease in Strength Due to Stranding. When a concentric-lay cable is subjected to mechanical tension, the spiral members tend to tighten around those layers under them and thus produce internal compression, gripping the inner layers and the core. Consequently, the individual wires, taken as a whole, do not behave as they would if they were true linear conductors acting independently. Furthermore, the individual wires are never exactly alike in diameter or in strength or in elastic properties. For these reasons, there is ordinarily a loss of about 4% to 11% in total tensile efficiency, depending on the number of layers. This reduction tends to increase as the pitch ratio decreases. Actual tensile tests on cables furnish the most dependable data on their ultimate strength.

Tensile efficiency of a stranded conductor is the ratio of its breaking strength to the sum of the tensile strengths of all its individual wires. Concentric-lay cables of 12 to 16 pitch ratio have a normal tensile efficiency of approximately 90%; rope-lay cables, approximately 80%.

Preformed Cable. This type of cable is made by preforming each individual wire (except the core) into a spiral of such length and curvature that the wire will fit naturally into its normal position in the cable instead of being forced into that shape under the usual tension in the stranding machine. This method has the advantage in cable made of the stiffer grades of wire that the individual wires do not tend to spread or untwist if the strand is cut in two without first binding the ends on each side of the cut.

Weight. A uniform cylindrical conductor of diameter d, length l, and density δ has a total weight expressed by the formula

$$W = \delta l \frac{\pi d^2}{4} \tag{4-16}$$

The weight of any conductor is commonly expressed in pounds per unit of length, such as 1 ft, 1000 ft, or 1 mi. The weight of stranded conductors can be calculated using Eq. (4-16), but allowance must be made for increase in weight due to stranding. Rope-lay stranding has greater increase in weight because of the multiple stranding operations.

Breaking Strength. The maximum load that a conductor attains when tested in tension to rupture.

Total Elongation at Rupture. When a sample of any material is tested under tension until it ruptures, measurement is usually made of the total elongation in a certain initial test length. In certain kinds of testing, the initial test length has been standardized, but in every case, the total elongation at rupture should be referred to the initial test length of the sample on which it was measured. Such elongation is usually expressed in percentage of original unstressed length and is a general index of the ductility of the material. Elongation is determined on solid conductors or on individual wires before stranding; it is rarely determined on stranded conductors.

Elasticity. All materials are deformed in greater or lesser degree under application of mechanical stress. Such deformation may be either of two kinds, known, respectively, as *elastic deformation* and *permanent deformation.* When a material is subjected to stress and undergoes deformation but resumes its original shape and dimensions when the stress is removed, the deformation is said to be *elastic*. If the stress is so great that the material fails to resume its original dimensions when the stress is removed, the permanent change in dimensions is termed *permanent deformation* or *set*. In general, the stress at which appreciable permanent deformation begins is termed the *working elastic limit*. Below this limit of stress the behavior of the material is said to be *elastic,* and in general, the deformation is proportional to the stress.

Stress and Strain. The *stress* in a material under load, as in simple tension or compression, is defined as the total load divided by the area of cross section normal to the direction of the load, assuming the load to be uniformly distributed over this cross section. It is commonly expressed in pounds per square inch. The *strain* in a material under load is defined as the total deformation measured in

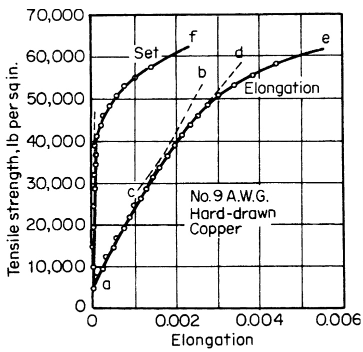


FIGURE 4-2 Stress-strain curves of No. 9 AWG hard-drawn copper wire. (Watertown Arsenal test).

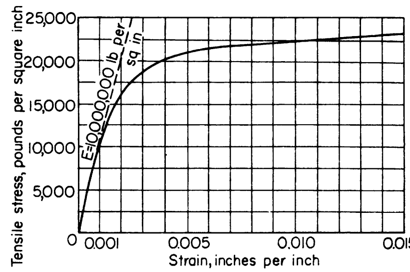


FIGURE 4-3 Typical stress-strain curve of hard drawn aluminum wire.

the direction of the stress, divided by the total unstressed length in which the measured deformation occurs, or the deformation per unit length. It is expressed as a decimal ratio or numeric.

In order to show the complete behavior of any given conductor under tension, it is customary to make a graph in terms of loading or stress as the ordinates and elongation or strain as the abscissas. Such graphs or curves are useful in determining the elastic limit and the yield point if the loading is carried to the point of rupture. Graphs showing the relationship between stress and strain in a material tested to failure are termed *load-deformation* or *stress-strain curves.*

Hooke's law consists of the simple statement that the stress is proportional to the strain. It obviously implies a condition of perfect elasticity, which is true only for stresses less than the elastic limit.

Stress-Strain Curves. A typical stress-strain diagram of hard-drawn copper wire is shown in Fig. 4-2, which represents No. 9 AWG. The curve *ae* is the actual stress-strain curve; *ab* represents the portion which corresponds to true elasticity, or for which Hooke's law holds rigorously; *cd* is the tangent *ae* which fixes the Johnson elastic limit; and the curve *af* represents the set, or permanent elongation due to flow of the metal under stress, being the difference between *ab* and *ae*. A typical stress-strain diagram of hard-drawn aluminum wire, based on data furnished by the Aluminum Company of America, is shown in Fig. 4-3.

Modulus (or Coefficient) of Elasticity. *Modulus (or coefficient) of elasticity* is the ratio of internal stress to the corresponding strain or deformation. It is a characteristic of each material, form (shape or structure), and type of stressing. For deformations involving changes in both volume and shape, special coefficients are used. For conductors under axial tension, the ratio of stress to strain is called *Young's modulus.*

If F is the total force or load acting uniformly on the cross section A, the stress is F/A. If this magnitude of stress causes an elongation e in an original length l, the strain is e/l. Young's modulus is then expressed

$$M = \frac{Fl}{Ae} \tag{4-17}$$

If a material were capable of sustaining an elastic elongation sufficient to make e equal to l, or such that the elongated length is double the original length, the stress required to produce this result would equal the modulus. This modulus is very useful in computing the sags of overhead conductor spans under loads of various kinds. It is usually expressed in pounds per square inch.

Stranding usually lowers the Young's modulus somewhat, rope-lay stranding to a greater extent than concentric-lay stranding. When a new cable is subjected initially to tension and the loading is carried up to the maximum working stress, there is an apparent elongation which is greater than the subsequent elongation under the same loading. This is apparently due to the removal of a very slight slackness in the individual wires, causing them to fit closely together and adjust themselves to the conditions of tension in the strand. When a new cable is loaded to the working limit, unloaded, and then reloaded, the value of Young's modulus determined on initial loading may be on the order of one-half to two-thirds of its true value on reloading. The latter figure should approach within a few percent of the modulus determined by test on individual straight wires of the same material.

For those applications where elastic stretching under tension needs consideration, the stress-strain curve should be determined by test, with the precaution not to prestress the cable before test unless it will be prestressed when installed in service. Commercially used values of Young's modulus for conductors are given in Table 4-1.

TABLE 4-1 Young's Moduli for Conductors

Conductor	Young's modulus,* lb/in^2		Reference
	Final†	Virtual initial‡	
Copper wire, hard-drawn	$17.0 \cdot 10^6$	$14.5 \cdot 10^6$	Copper Wire Engineering Assoc.
Copper wire, medium hard-drawn	$16.0 \cdot 10^6$	$14.0 \cdot 10^6$	Anaconda Wire and Cable Co.
Copper cable, hard-drawn, 3 and 12 wire	$17.0 \cdot 10^6$	$14.0 \cdot 10^6$	Copper Wire Engineering Assoc.
Copper cable, hard-drawn, 7 and 19 wire	$17.0 \cdot 10^6$	$14.5 \cdot 10^6$	Copper Wire Engineering Assoc.
Copper cable, medium hard-drawn	$15.5 \cdot 10^6$	$14.0 \cdot 10^6$	Anaconda Wire and Cable Co.
Bronze wire, alloy 15	$14.0 \cdot 10^6$	$13.0 \cdot 10^6$	Anaconda Wire and Cable Co.
Bronze wire, other alloys	$16.0 \cdot 10^6$	$14.0 \cdot 10^6$	Anaconda Wire and Cable Co.
Bronze cable, alloy 15	$13.0 \cdot 10^6$	$12.0 \cdot 10^6$	Anaconda Wire and Cable Co.
Bronze cable, other alloys	$16.0 \cdot 10^6$	$14.0 \cdot 10^6$	Anaconda Wire and Cable Co.
Copper-clad steel wire	$24.0 \cdot 10^6$	$22.0 \cdot 10^6$	Copperweld Steel Co.
Copper-clad steel cable	$23.0 \cdot 10^6$	$20.5 \cdot 10^6$	Copperweld Steel Co.
Copper–copper-clad steel cable, type E	$19.5 \cdot 10^6$	$17.0 \cdot 10^6$	Copperweld Steel Co.
Copper–copper-clad steel cable, type EK	$18.5 \cdot 10^6$	$16.0 \cdot 10^6$	Copperweld Steel Co.
Copper–copper-clad steel cable, type F	$18.0 \cdot 10^6$	$15.5 \cdot 10^6$	Copperweld Steel Co.
Copper–copper-clad steel cable, type 2A to 6A	$19.0 \cdot 10^6$	$16.5 \cdot 10^6$	Copper Wire Engineering Assoc.
Aluminum wire	$10.0 \cdot 10^6$		Reynolds Metals Co.
Aluminum cable	$9.1 \cdot 10^6$	$7.3 \cdot 10^6$	Reynolds Metals Co.
Aluminum-alloy wire	$10.0 \cdot 10^6$		Reynolds Metals Co.
Aluminum-alloy cable	$9.1 \cdot 10^6$	$7.3 \cdot 10^6$	Reynolds Metals Co.
Aluminum-steel cable, aluminum wire	$7.2–9.0 \cdot 10^6$		Aluminum Co. of America
Aluminum-steel cable, steel wire	$26.0–29.0 \cdot 10^6$		Aluminum Co. of America
Aluminum-clad steel wire	$23.5 \cdot 10^6$	$22.0 \cdot 10^6$	Copperweld Steel Co.
Aluminum-clad steel cable	$23.0 \cdot 10^6$	$21.5 \cdot 10^6$	Copperweld Steel Co.
Aluminum-clad steel–aluminum cable:			
AWAC 5/2	$13.5 \cdot 10^6$	$12.0 \cdot 10^6$	Copperweld Steel Co.
AWAC 4/3	$15.5 \cdot 10^6$	$14.0 \cdot 10^6$	Copperweld Steel Co.
AWAC 3/4	$17.5 \cdot 10^6$	$16.0 \cdot 10^6$	Copperweld Steel Co.
AWAC 2/5	$19.0 \cdot 10^6$	$18.0 \cdot 10^6$	Copperweld Steel Co.
Galvanized-steel wire, Class A coating	$28.5 \cdot 10^6$		Indiana Steel & Wire Co.
Galvanized-steel cable, Class A coating	$27.0 \cdot 10^6$		Indiana Steel & Wire Co.

Note: 1 lb/in^2 = 6.895 kPa.

*For stranded cables the moduli are usually less than for solid wire and vary with number and arrangement of strands, tightness of stranding, and length of lay. Also, during initial application of stress, the stress-strain relation follows a curve throughout the upper part of the range of stress commonly used in transmission-line design.

†Final modulus is the ratio of stress to strain (slope of the curve) obtained after fully prestressing the conductor. It is used in calculating design or final sags and tensions.

‡Virtual initial modulus is the ratio of stress to strain (slope of the curve) obtained during initial sustained loading of new conductor. It is used in calculating initial or stringing sags and tensions.

Young's Modulus for ACSR. The permanent modulus of ACSR depends on the proportions of steel and aluminum in the cable and on the distribution of stress between aluminum and steel. This latter condition depends on temperature, tension, and previous maximum loadings. Because of the interchange of stress between the steel and the aluminum caused by changes of tension and temperature, computer programs are ordinarily used for sag-tension calculations.

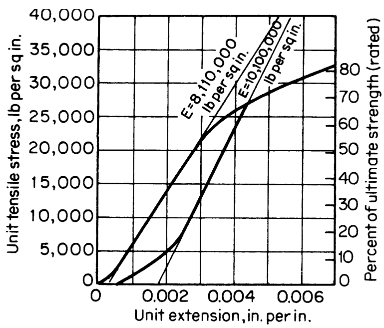


FIGURE 4-4 Repeated stress-strain curve, 795,000 cmil ACSR; 54 × 0.1212 aluminum strands, 7 × 0.1212 steel strands.

Because ACSR is a composite cable made of aluminum and steel wires, additional phenomena occur which are not found in tests of cable composed of a single material. As shown in Fig. 4-4, the part of the curve obtained in the second stress cycle contains a comparatively large "foot" at its base, which is caused by the difference in extension at the elastic limits of the aluminum and steel.

Elastic Limit. This is variously defined as the limit of stress beyond which permanent deformation occurs or the stress limit beyond which Hooke's law ceases to apply or the limit beyond which the stresses are not proportional to the strains or the *proportional limit*. In some materials, the elastic limit occurs at a point which is readily determined, but in others it is quite difficult to determine because the stress-strain curve deviates from a straight line but very slightly at first, and the point of departure from true linear relationship between stress and strain is somewhat indeterminate.

Dean J. B. Johnson of the University of Wisconsin, well-known authority on materials of construction, proposed the use of an arbitrary determination referred to frequently as the *Johnson definition of elastic limit*. This proposal, which has been quite largely used, was that an *apparent elastic limit* be employed, defined as that point on the stress-strain curve at which the rate of deformation is 50% greater than at the origin. The apparent elastic limit thus defined is a practical value, which is suitable for engineering purposes because it involves negligible permanent elongation.

The *Johnson elastic limit* is that point on the stress-strain curve at which the natural tangent is equal to 1.5 times the tangent of the angle of the straight or linear portion of the curve, with respect to the axis of ordinates, or *Y* axis.

Yield Point. In many materials, a point is reached on the stress-strain diagram at which there is a marked increase in strain or elongation without an increase in stress or load. The point at which this occurs is termed the *yield point*. It is usually quite noticeable in ductile materials but may be scarcely perceptible or possibly not present at all in certain hard-drawn materials such as hard-drawn copper.

Prestressed Conductors. In the case of some materials, especially those of considerable ductility, which tend to show permanent elongation or "drawing" under loads just above the initial elastic limit, it is possible to raise the working elastic limit by loading them to stresses somewhat above the elastic limit as found on initial loading. After such loading, or prestressing, the material will behave according to Hooke's law at all loads less than the new elastic limit. This applies not only to many ductile materials, such as soft or annealed copper wire, but also to cables or stranded conductors, in which there is a slight inherent slack or looseness of the individual wires that can be removed only under actual loading. It is sometimes the practice, when erecting such conductors for service, to prestress them to the working elastic limit or safe maximum working stress and then reduce the stress to the proper value for installation at the stringing temperature without wind or ice.

Resistance. *Resistance* is the property of an electric circuit or of any body that may be used as part of an electric circuit which determines for a given current the average rate at which electrical energy is converted into heat. The term is properly applied only when the rate of conversion is proportional to the square of the current and is then equal to the power conversion divided by the square of the current. A uniform cylindrical conductor of diameter d, length l, and *volume resistivity* ρ has a total resistance to continuous currents expressed by the formula

$$R = \frac{\rho l}{\pi d^2/4} \tag{4-18}$$

The resistance of any conductor is commonly expressed in ohms per unit of length, such as 1 ft, 1000 ft, or 1 mi. When used for conducting alternating currents, the effective resistance may be higher than the dc resistance defined above. In the latter case, it is a common practice to apply the proper factor, or ratio of effective ac resistance to dc resistance, sometimes termed the *skin-effect resistance ratio*. This ratio may be determined by test, or it may be calculated if the necessary data are available.

Magnetic Permeability. *Magnetic permeability* applies to a field in which the flux is uniformly distributed over a cross section normal to its direction or to a sufficiently small cross section of a nonuniform field so that the distribution can be assumed as substantially uniform. In the case of a cylindrical conductor, the magnetomotive force (mmf) due to the current flowing in the conductor varies from zero at the center or axis to a maximum at the periphery or surface of the conductor and sets up a flux in circular paths concentric with the axis and perpendicular to it but of nonuniform distribution between the axis and the periphery. If the permeability is nonlinear with respect to the mmf, as is usually true with magnetic materials, there is no correct single value of permeability which fits the conditions, although an apparent or equivalent average value can be determined. In the case of other forms of cross section, the distribution is still more complex, and the equivalent permeability may be difficult or impossible to determine except by test.

Internal Inductance. A uniform cylindrical conductor of nonmagnetic material, or of unit permeability, has a constant magnitude of internal inductance per unit length, independent of the conductor diameter. This is commonly expressed in microhenrys or millihenrys per unit of length, such as 1 ft, 1000 ft, or 1 mi. When the conductor material possesses magnetic susceptibility, and when the magnetic permeability μ is constant and therefore independent of the current strength, the internal inductance is expressed in absolute units by the formula

$$L = \frac{\mu l}{2} \tag{4-19}$$

In most cases, μ is not constant but is a function of the current strength. When this is true, there is an effective permeability, one-half of which ($\mu/2$) expresses the inductance per centimeter of length, but this figure of permeability is virtually the ratio of the effective inductance of the conductor of susceptible material to the inductance of a conductor of material which has a permeability of unity. When used for conducting alternating currents, the effective inductance may be less than the inductance with direct current; this is also a direct consequence of the same skin effect which results in an increase of effective resistance with alternating currents, but the overall effect is usually included in the figure of effective permeability. It is usually the practice to determine the effective internal inductance by test, but it may be calculated if the necessary data are available.

Skin Effect. *Skin effect* is a phenomenon which occurs in conductors carrying currents whose intensity varies rapidly from instant to instant but does not occur with continuous currents. It arises from the fact that elements or filaments of variable current at different points in the cross section of a conductor do not encounter equal components of inductance, but the central or axial filament meets the maximum inductance, and in general the inductance offered to other filaments of current decreases as the distance of the filament from the axis increases, becoming a minimum at the surface or periphery of the conductor. This, in turn, tends to produce unequal current density over the cross section as a whole; the density is a minimum at the axis and a maximum at the periphery. Such distribution of the current density produces an increase in effective resistance and a decrease in effective internal inductance; the former is of more practical importance than the latter. In the case of large copper conductors at commercial power frequencies and in the case of most conductors at carrier and radio frequencies, the increase in resistance should be considered.

Skin-Effect Ratios. If R' is the effective resistance of a linear cylindrical conductor to sinusoidal alternating current of given frequency and R is the true resistance with continuous current, then

$$R' = KR \quad \text{ohms} \tag{4-20}$$

where K is determined from Table 4-2 in terms of x. The value of x is given by

$$x = 2\pi a \sqrt{\frac{2f\mu}{\rho}} \tag{4-21}$$

where a is the radius of the conductor in centimeters, f is the frequency in cycles per second, μ is the magnetic permeability of the conductor (here assumed to be constant), and ρ is the resistivity in abohm-centimeters (abohm $= 10^{-9}\ \Omega$).

TABLE 4-2 Skin-Effect Ratios

x	K	K'	x	K	K'	x	K	K'	x	K	K'
0.0	1.00000	1.00000	2.9	1.28644	0.86012	6.6	2.60313	0.42389	17.0	6.26817	0.16614
0.1	1.00000	1.00000	3.0	1.31809	0.84517	6.8	2.67312	0.41171	18.0	6.62129	0.15694
0.2	1.00001	1.00000	3.1	1.35102	0.82975	7.0	2.74319	0.40021	19.0	6.97446	0.14870
0.3	1.00004	0.99998	3.2	1.38504	0.81397	7.2	2.81334	0.38933	20.0	7.32767	0.14128
0.4	1.00013	0.99993	3.3	1.41999	0.79794	7.4	2.88355	0.37902	21.0	7.68091	0.13456
0.5	1.00032	0.99984	3.4	1.45570	0.78175	7.6	2.95380	0.36923	22.0	8.03418	0.12846
0.6	1.00067	0.99966	3.5	1.49202	0.76550	7.8	3.02411	0.35992	23.0	8.38748	0.12288
0.7	1.00124	0.99937	3.6	1.52879	0.74929	8.0	3.09445	0.35107	24.0	8.74079	0.11777
0.8	1.00212	0.99894	3.7	1.56587	0.73320	8.2	3.16480	0.34263	25.0	9.09412	0.11307
0.9	1.00340	0.99830	3.8	1.60314	0.71729	8.4	3.23518	0.33460	26.0	9.44748	0.10872
1.0	1.00519	0.99741	3.9	1.64051	0.70165	8.6	3.30557	0.32692	28.0	10.15422	0.10096
1.1	1.00758	0.99621	4.0	1.67787	0.68632	8.8	3.37597	0.31958	30.0	10.86101	0.09424
1.2	1.01071	0.99465	4.1	1.71516	0.67135	9.0	3.44638	0.31257	32.0	11.56785	0.08835
1.3	1.01470	0.99266	4.2	1.75233	0.65677	9.2	3.51680	0.30585	34.0	12.27471	0.08316
1.4	1.01969	0.99017	4.3	1.78933	0.64262	9.4	3.58723	0.29941	36.0	12.98160	0.07854
1.5	1.02582	0.98711	4.4	1.82614	0.62890	9.6	3.65766	0.29324	38.0	13.68852	0.07441
1.6	1.03323	0.98342	4.5	1.86275	0.61563	9.8	3.72812	0.28731	40.0	14.39545	0.07069
1.7	1.04205	0.97904	4.6	1.89914	0.60281	10.0	3.79857	0.28162	42.0	15.10240	0.06733
1.8	1.05240	0.97390	4.7	1.93533	0.59044	10.5	3.97477	0.26832	44.0	15.80936	0.06427
1.9	1.06440	0.96795	4.8	1.97131	0.57852	11.0	4.15100	0.25622	46.0	16.51634	0.06148
2.0	1.07816	0.96113	4.9	2.00710	0.56703	11.5	4.32727	0.24516	48.0	17.22333	0.05892
2.1	1.09375	0.95343	5.0	2.04272	0.55597	12.0	4.50358	0.23501	50.0	17.93032	0.05656
2.2	1.11126	0.94482	5.2	2.11353	0.53506	12.5	4.67993	0.22567	60.0	21.46541	0.04713
2.3	1.13069	0.93527	5.4	2.18389	0.51566	13.0	4.85631	0.21703	70.0	25.00063	0.04040
2.4	1.15207	0.92482	5.6	2.25393	0.49764	13.5	5.03272	0.20903	80.0	28.53593	0.03535
2.5	1.17538	0.91347	5.8	2.32380	0.48086	14.0	5.20915	0.20160	90.0	32.07127	0.03142
2.6	1.20056	0.90126	6.0	2.39359	0.46521	14.5	5.38560	0.19468	100.0	35.60666	0.02828
2.7	1.22753	0.88825	6.2	2.46338	0.45056	15.0	5.56208	0.18822	∞	∞	0
2.8	1.25620	0.87451	6.4	2.53321	0.43682	16.0	5.91509	0.17649			

For practical calculation, Eq. (4-21) can be written

$$x = 0.063598\sqrt{\frac{f\mu}{R}} \tag{4-22}$$

where R is dc resistance at operating temperature in ohms per mile.

If L' is the effective inductance of a linear conductor to sinusoidal alternating current of a given frequency, then

$$L' = L_1 + K'L_2 \tag{4-23}$$

where L_1 is external portion of inductance, L_2 is internal portion (due to the magnetic field within the conductor), and K' is determined from Table 4-2 in terms of x. Thus, the total effective inductance per unit length of conductor is

$$L' = 2\ln\frac{d}{a} + K'\frac{\mu}{2} \tag{4-24}$$

The inductance is here expressed in abhenrys per centimeter of conductor, in a linear circuit; a is the radius of the conductor, and d is the separation between the conductor and its return conductor, expressed in the same units.

Values of K and K' in terms of x are shown in Table 4-2 and Figs. 4-5 and 4-6 (see *NBS Circ*. 74, pp. 309–311, for additional tables, and *Sci. Paper* 374).

Value of μ for nonmagnetic materials (copper, aluminum, etc.) is 1; for magnetic materials, it varies widely with composition, processing, current density, etc., and should be determined by test in each case.

Alternating-Current Resistance. For small conductors at power frequencies, the frequency has a negligible effect, and dc resistance values can be used. For large conductors, frequency must be taken into account in addition to temperature effects. To do this, first calculate the dc resistance at the operating temperature, then determine the skin-effect ratio K, and finally determine the ac resistance at operating temperature.

AC resistance for copper conductors *not* in close proximity can be obtained from the skin-effect ratios given in Tables 4-2 and 4-3.

AC Resistance for Aluminum Conductors. The increase in resistance and decrease in internal inductance of cylindrical aluminum conductors can be determined from data. It is not the same as for copper conductors of equal diameter but is slightly less because of the higher volume resistivity of aluminum.

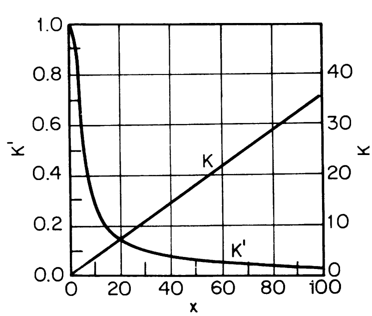


FIGURE 4-5 K and K' for values of x from 0 to 100.

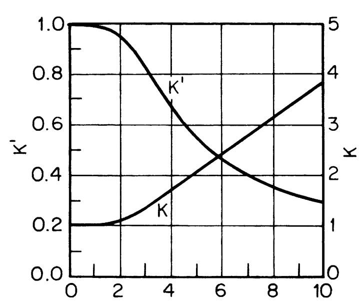


FIGURE 4-6 K and K' for values of x from 0 to 10.

TABLE 4-3 Skin-Effect Ratios—Copper Conductors *Not* in Close Proximity

Conductor size, Mcm	Skin-effect ratio ***K*** at 60 cycles and 65°C (149°F)															
	Inside conductor diameter, in															
	0*		0.25		0.50		0.75		1.00		1.25		1.30		2.00	
	Outside diameter, in	***K***	Outside diameter, in	***K***	Outside diameter, in	***K***	Outside diameter, in	***K***	Outside diameter, in	***K***	Outside diameter, in	***K***	Outside diameter, in	***K***	Outside diameter, in	***K***
3000	1.998	1.439	2.02	1.39	2.08	1.36	2.15	1.29	2.27	1.23	2.39	1.19	2.54	1.15	2.87	1.08
2500	1.825	1.336	1.87	1.28	1.91	1.24	2.00	1.20	2.12	1.16	2.25	1.12	2.40	1.09	2.75	1.05
2000	1.631	1.239	1.67	1.20	1.72	1.17	1.80	1.12	1.94	1.09	2.09	1.06	2.25	1.05	2.61	1.02
1500	1.412	1.145	1.45	1.12	1.52	1.09	1.63	1.06	1.75	1.04	1.91	1.03	2.07	1.02	2.47	1.01
1000	1.152	1.068	1.19	1.05	1.25	1.03	1.39	1.02	1.53	1.01	1.72	1.01				
800	1.031	1.046	1.07	1.04	1.16	1.02	1.28	1.01	1.45	1.01						
600	0.893	1.026	0.94	1.02	1.04	1.01										
500	0.814	1.018	0.86	1.01	0.97	1.01										
400	0.728	1.012	0.78	1.01												
300	0.630	1.006														

Note: 1 in = 2.54 cm.
*For standard concentric-stranded conductors (i.e., inside diameter = 0).

AC Resistance for ACSR. In the case of ACSR conductors, the steel core is of relatively high resistivity, and therefore its conductance is usually neglected in computing the total resistance of such strands. The effective permeability of the grade of steel employed in the core is also relatively small. It is approximately correct to assume that such a strand is hollow and consists exclusively of its aluminum wires; in this case, the laws of skin effect in tubular conductors will be applicable. Conductors having a single layer of aluminum wires over the steel core have higher ac/dc ratios than those having multiple layers of aluminum wires.

Inductive Reactance. Present practice is to consider inductive reactance as split into two components: (1) that due to flux within a radius of 1 ft including the internal reactance within the conductor of radius r and (2) that due to flux between 1 ft radius and the equivalent conductor spacing D_s or geometric mean distance (GMD). The fundamental inductance formula is

$$L = 2 \ln \frac{D_s}{r} + \frac{\mu}{2} \qquad \text{abH/(cm)(conductor)} \tag{4-25}$$

This can be rewritten

$$L = 2 \ln \frac{D_s}{1} + 2 \ln \frac{1}{r} + \frac{\mu}{2} \tag{4-26}$$

where the term $2 \ln (D_s/1)$ represents inductance due to flux between 1 ft radius and the equivalent conductor spacing, and $2 \ln (1/r) + (\mu/2)$ represents the inductance due to flux within 1 ft radius [$2 \ln (1/r)$ represents inductance due to flux between conductor surface and 1 ft radius, and $\mu/2$ represents internal inductance due to flux within the conductor].

By definition, *geometric mean radius* (GMR) of a conductor is the radius of an infinitely thin tube having the same internal inductance as the conductor. Therefore,

$$L = 2 \ln \frac{D_s}{1} + 2 \ln \frac{1}{\text{GMR}} \tag{4-27}$$

Since inductive reactance $= 2\pi fL$, for practical calculation Eq. (4-27) can be written

$$X = 0.004657 f \log \frac{D_s}{1} + 0.004657 f \log \frac{1}{\text{GMR}} \qquad \Omega\text{/(mi)(conductor)} \tag{4-28}$$

In the conductor tables in this section, inductive reactance is calculated from Eq. (4-28), considering that

$$X = x_a + x_d \tag{4-29}$$

Inductive reactance for conductors using steel varies in a manner similar to ac resistance.

Capacitive Reactance. The capacitive reactance can be considered in two parts also, giving

$$X = \frac{4.099}{f} \log \frac{D_s}{1} + \frac{4.099}{f} \log \frac{1}{r} \qquad \text{M}\Omega\text{/(mi)(conductor)} \tag{4-30}$$

In the conductor tables in this section, capacitive reactance is calculated from Eq. (4-30), it being considered that

$$X' = x'_a + x'_d \tag{4-31}$$

It is important to note that in capacitance calculations the conductor radius used is the actual physical radius of the conductor.

Capacitive Susceptance

$$B = \frac{1}{x_a' + x_d'} \qquad \mu\text{S(mi)(conductor)} \tag{4-32}$$

Charging Current

$$I_C = eB \times 10^{-3} \qquad \text{A/(mi)(conductor)} \tag{4-33}$$

where e is voltage to neutral in kilovolts.

Bus Conductors. *Bus conductors* require that greater attention be given to certain physical and electrical characteristics of the metals than is usually necessary in designing line conductors. These characteristics are current-carrying capacity, emissivity, skin effect, expansion, and mechanical deflection. To obtain the most satisfactory and economical designs for bus bars in power stations and substations, where they are used extensively, consideration must be given to choice not only of material but also of shape. Both copper and aluminum are used for bus bars, and in certain outdoor substations, steel has proved satisfactory. The most common bus bar form for carrying heavy current, especially indoors, is flat copper bar. Bus bars in the form of angles, channels, and tubing have been developed for heavy currents and, because of better distribution of the conducting material, make more efficient use of the metal both electrically and mechanically. All such designs are based on the need for proper current-carrying capacity without excess bus bar temperatures and on the necessity for adequate mechanical strength.

Hollow (Expanded) Conductors. Hollow (expanded) conductors are used on high-voltage transmission lines when, in order to reduce corona loss, it is desirable to increase the outside diameter without increasing the area beyond that needed for maximum line economy. Not only is the initial corona voltage considerably higher than for conventional conductors of equal cross section, but the current-carrying capacity for a given temperature rise is also greater because of the larger surface area available for cooling and the better disposition of the metal with respect to skin effect when carrying alternating currents.

Air-expanded ACSR is a conductor whose diameter has been increased by aluminum skeletal wires between the steel core and the outer layers of aluminum strands creating air spaces. A conductor having the necessary diameter to minimize corona effects on lines operating above 300 kV will, many times, have more metal than is economical if the conductor is made conventionally.

Composite Conductors. *Composite conductors* are those made up of usually two different types of wire having differing characteristics. They are generally designed for a ratio of physical and electrical characteristics different from those found in homogeneous materials. *Aluminum conductors, steel reinforced (ACSR)* and *aluminum conductors, aluminum alloy reinforced (ACAR)* are types commonly used in overhead transmission and distribution lines.

Cables of this type are particularly adaptable to long-span construction or other service conditions requiring more than average strength combined with liberal conductance. They lend themselves readily to economical, dependable use on transmission lines, rural distribution lines, railroad electrification, river crossings, and many kinds of special construction.

Self-damping ACSR conductors are used to limit aeolian vibration to a safe level regardless of conductor tension or span length. They are concentrically stranded conductors composed of two layers of trapezoidal-shaped wires or two layers of trapezoidal-shaped wires and one layer of round wires of 1350 (EC) alloy with a high-strength, coated steel core. The trapezoidal wire layers are self-supporting, and separated by gaps from adjacent layers (Fig. 4-7). Impact between layers during aeolian vibration causes damping action.

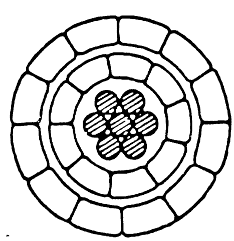


FIGURE 4-7 Self-damping ACSR conductor.

ACSR/TW is similar to self-damping ACSR in its use of trapezoidal-shaped wires, but does not have the annular gaps between layers. ACSR/TW has a smaller diameter and smoother surface than conventional round-wire ACSR of the same area, and thus may have reduced wind loading.

T2 conductors are fabricated by twisting two conventional conductors together with a pitch of about 9 ft (2.7 m). Severity of wind-induced galloping when the conductor is coated with ice is reduced because an ice profile that is uniform along the conductor length cannot form on the variable profile presented by the conductor.

Steel-supported aluminum conductors (SSAC) are similar to conventional ACSR but employ an aluminum alloy in the annealed condition. The annealed aluminum has increased electrical conductivity, and the conductor has improved sag-tension characteristics for high-temperature service.

4.1.4 Fusible Metals and Alloys

Fusible alloys having melting points in the range from about 60 to 200°C are made principally of bismuth, cadmium, lead, and tin in various proportions. Many of these alloys have been known under the names of their inventors (see index of alloys in *International Critical Tables*, vol. 2).

Fuse metals for electric fuses of the open-link enclosed and expulsion types are ordinarily made of some low-fusible alloy; aluminum also is used to some extent. The resistance of the fuse causes dissipation of energy, liberation of heat, and rise of temperature. Sufficient current obviously will melt the fuse, and thus open the circuit if the resulting arc is self-extinguishing. Metals which volatilize readily in the heat of the arc are to be preferred to those which leave a residue of globules of hot metal. The rating of any fuse depends critically on its shape, dimensions, mounting, enclosure, and any other factors which affect its heat-dissipating capacity.

Fusing currents of different kinds of wire were investigated by W. H. Preece, who developed the formula

$$I = ad^{3/2} \tag{4-34}$$

where I is fusing current in amperes, d is diameter of the wire in inches, and a is a constant depending on the material. He found the following values for a:

Copper	10,244	Iron	3,148
Aluminum	7,585	Tin	1,642
Platinum	5,172	Alloy (2Pb-1Sn)	1,318
German silver	5,230	Lead	1,379
Platinoid	4,750		

Although this formula has been used to a considerable extent in the past, it gives values that usually are erroneous in practice because it is based on the assumption that all heat loss is due to radiation. A formula of the general type

$$I = kd^n \tag{4-35}$$

can be used with accuracy if k and n are known for the particular case (material, wire size, installation conditions, etc.).

Fusing current-time for copper conductors and connections may be determined by an equation developed by I. M. Onderdonk

$$33\left(\frac{I}{A}\right)^2 S = \log\left(\frac{T_m - T_a}{234 + T_a} + 1\right) \tag{4-36}$$

$$I = A\sqrt{\frac{\log\left(\frac{T_m - T_a}{234 + T_a} + 1\right)}{33S}} \tag{4-37}$$

where I is current in amperes, A is conductor area in circular mils, S is time current applied in seconds, T_m is melting point of copper in degrees Celsius, and T_a is ambient temperature in degrees Celsius.

4.1.5 Miscellaneous Metals and Alloys

Contact Metals. *Contact metals* may be grouped into three general classifications:

Hard metals, which have high melting points, for example, tungsten and molybdenum. Contacts of these metals are employed usually where operations are continuous or very frequent and current has nominal value of 5 to 10 A. Hardness to withstand mechanical wear and high melting point to resist arc erosion and welding are their outstanding advantages. A tendency to form high-resistance oxides is a disadvantage, but this can be overcome by several methods, such as using high-contact force, a hammering or wiping action, and a properly balanced electric circuit.

Highly conductive metals, of which silver is the best for both electric current and heat. Its disadvantages are softness and a tendency to pit and transfer. In sulfurous atmosphere, a resistant sulfide surface will form on silver, which results in high contact-surface resistance. These disadvantages are overcome usually by alloying.

Noncorroding metals, which for the most part consist of the noble metals, such as gold and the platinum group. Contacts of these metals are used on sensitive devices, employing extremely light pressures or low current in which clean contact surfaces are essential. Because most of these metals are soft, they are usually alloyed.

The metals commonly used are tungsten, molybdenum, platinum, palladium, gold, silver, and their alloys. Alloying materials are copper, nickel, cadmium, iron, and the rarer metals such as iridium and ruthenium. Some are prepared by powder metallurgy.

Tungsten. *Tungsten* (W) is a hard, dense, slow-wearing metal, a good thermal and electrical conductor, characterized by its high melting point and freedom from sticking or welding. It is manufactured in several grades having various grain sizes.

Molybdenum. *Molybdenum* (Mo) has contact characteristics about midway between tungsten and fine silver. It often replaces either metal where greater wear resistance than that of silver or lower contact-surface resistance than that of tungsten is desired.

Platinum. *Platinum* (Pt) is one of the most stable of all metals under the combined action of corrosion and electrical erosion. It has a high melting point and does not corrode and surfaces remain clean and low in resistance under most adverse atmospheric and electrical conditions. *Platinum alloys* of iridium (Ir), ruthenium (Ru), silver (Ag), or other metals are used to increase hardness and resistance to wear.

Palladium. *Palladium* (Pd) has many of the properties of platinum and is frequently used as an alternate for platinum and its alloys. *Palladium alloys* of silver (Ag), ruthenium (Ru), nickel (Ni), and other metals are used to increase hardness and resistance to wear.

Gold. *Gold* (Au) is similar to platinum in corrosion resistance but has a much lower melting point. Gold and its alloys are ductile and easily formed into a variety of shapes. Because of its softness, it is usually alloyed. *Gold alloys* of silver (Ag) and other metals are used to impart hardness and improve resistance to mechanical wear and electrical erosion.

Silver. *Silver* (Ag) has the highest thermal and electrical conductivity (110%, IACS) of any metal. It has low contact-surface resistance, since its oxide decomposes at approximately 300°F. It is available commercially in three grades:

Grade	Typical composition, %	
	Silver	Copper
Fine silver	99.95+	
Sterling silver	92.5	7.5
Coin silver	90	10

Fine silver is used extensively under low contact pressure where sensitivity and low contact-surface resistance are essential or where the circuit is operated infrequently.

Sterling and coin silvers are harder than fine silver and resist transfer at low voltage (6 to 8 V) better than fine silver. Since their contact-surface resistance is greater than that of fine silver, higher contact-closing forces should be used.

Silver alloys of copper (Cu), nickel (Ni), cadmium (Cd), iron (Fe), carbon (C), tungsten (W), molybdenum (Mo), and other metals are used to improve hardness, resistance to wear and arc erosion, and for special applications.

Selenium. *Selenium* is a nonmetallic element chemically resembling sulfur and tellurium and occurs in several allotropic forms varying in specific gravity from 4.3 to 4.8. It melts at 217°C and boils at 690°C. At 0°C, it has a resistivity of approximately 60,000 Ω · cm. The dielectric constant ranges from 6.1 to 7.4. It has the peculiar property that its resistivity decreases on exposure to light; the resistivity in darkness may be anywhere from 5 to 200 times the resistivity under exposure to light.

4.2 MAGNETIC MATERIALS

4.2.1 Definitions

The following definitions of terms relating to magnetic materials and the properties and testing of these materials have been selected from ASTM Standard. Terms primarily related to magnetostatics are indicated by the symbol * and those related to magnetodynamics are indicated by the symbol **. General (nonrestricted) terms are not marked.

***AC Excitation N_1I/l_1.* The ratio of the rms ampere-turns of exciting current in the primary winding of an inductor to the effective length of the magnetic path.

***Active (Real) Power P.* The product of the rms current I in an electric circuit, the rms voltage E across the circuit, and the cosine of the angular phase difference θ between the current and the voltage.

$$P = EI\cos\theta \tag{4-38}$$

NOTE: The portion of the active power that is expended in a magnetic core is the total core loss P_c.

Aging, Magnetic. The change in the magnetic properties of a material resulting from metallurgical change. This term applies whether the change results from a continued normal or a specified accelerated aging condition.

NOTE: This term implies a deterioration of the magnetic properties of magnetic materials for electronic and electrical applications, unless otherwise specified.

Ampere-turn. Unit of magnetomotive force in the rationalized mksa system. One ampere-turn equals 4π/10, or 1.257 gilberts.

Ampere-turn per Meter. Unit of magnetizing force (magnetic field strength) in the rationalized mksa system. One ampere-turn per meter is $4\pi \times 10^{-3}$, or 0.01257 oersted.

Anisotropic Material. A material in which the magnetic properties differ in various directions.

Antiferromagnetic Material. A feebly magnetic material in which almost equal magnetic moments are lined up antiparallel to each other. Its susceptibility increases as the temperature is raised until a critical (Neél) temperature is reached; above this temperature the material becomes paramagnetic.

**Apparent Power P_a.* The product (volt-amperes) of the rms exciting current and the applied rms *terminal* voltage in an *electric* circuit containing inductive impedance. The components of this impedance due to the winding will be linear, while the components due to the magnetic core will be nonlinear.

**Apparent Power; Specific, $P_{a(B,f)}$.* The value of the apparent power divided by the active mass of the specimen (volt-amperes per unit mass) taken at a specified maximum value of cyclically varying induction B and at a specified frequency f.

Coercive Force H_c. The (dc) magnetizing force at which the magnetic induction is zero when the material is in a symmetrically cyclically magnetized condition.

Coercive Force, Intrinsic, H_{ci}. The (dc) magnetizing force at which the intrinsic induction is zero when the material is in a symmetrically cyclically magnetized condition.

Coercivity H_{cs}. The maximum value of coercive force.

**Core Loss; Specific, $P_{c(B,f)}$.* The active power (watts) expended per unit mass of magnetic material in which there is a cyclically varying induction of a specified maximum value B at a specified frequency f.

**Core Loss (Total) P_c.* The active power (watts) expended in a magnetic circuit in which there is a cyclically alternating induction.

NOTE: Measurements of core loss are normally made with sinusoidally alternating induction, or the results are corrected for deviations from the sinusoidal condition.

Curie Temperature T_c. The temperature above which a ferromagnetic material becomes paramagnetic.

Demagnetization Curve. That portion of a normal (dc) hysteresis loop which lies in the second or fourth quadrant, that is, between the residual induction point B_r and the coercive force point H_c. Points on this curve are designated by the coordinates B_d and H_d.

Diamagnetic Material. A material whose relative permeability is less than unity.

NOTE: The intrinsic induction B_i, is oppositely directed to the applied magnetizing force H.

Domains, Ferromagnetic. Magnetized regions, either macroscopic or microscopic in size, within ferromagnetic materials. Each domain per se is magnetized to intrinsic saturation at all times, and this saturation induction is unidirectional within the domain.

**Eddy-Current Loss, Normal, P_e.* That portion of the core loss which is due to induced currents circulating in the magnetic material subject to an *SCM* excitation.

Energy Product B_dH_d. The product of the coordinate values of any point on a demagnetization curve.

Energy-Product Curve, Magnetic. The curve obtained by plotting the product of the corresponding coordinates B_d and H_d of points on the demagnetization curve as abscissa against the induction B_d as ordinates.

NOTE 1: The maximum value of the energy product $(B_dH_d)_m$ corresponds to the maximum value of the external energy.

NOTE 2: The demagnetization curve is plotted to the left of the vertical axis and usually the energy-product curve to the right.

**Exciting Power, rms, P_z.* The product of the rms exciting current and the rms voltage induced in the exciting (primary) winding on a magnetic core.

NOTE: This is the apparent volt-amperes required for the excitation of the magnetic core only. When the core has a secondary winding, the induced primary voltage is obtained from the measured open-circuit secondary voltage multiplied by the appropriate turns ratio.

**Exciting Power, Specific* $P_{z(B,f)}$. The value of the rms exciting power divided by the active mass of the specimen (volt-amperes/unit mass) taken at a specified maximum value of cyclically varying induction B and at specified frequency f.

Ferrimagnetic Material. A material in which unequal magnetic moments are lined up antiparallel to each other. Permeabilities are of the same order of magnitude as those of ferromagnetic materials, but are lower than they would be if all atomic moments were parallel and in the same direction. Under ordinary conditions, the magnetic characteristics of ferrimagnetic materials are quite similar to those of ferromagnetic materials.

Ferromagnetic Material. A material that, in general, exhibits the phenomena of hysteresis and saturation, and whose permeability is dependent on the magnetizing force.

Gauss (Plural Gausses). The unit of magnetic induction in the cgs electromagnetic system. The gauss is equal to 1 maxwell per square centimeter or 10^{-4} T. See *magnetic induction (flux density)*.

Gilbert. The unit of magnetomotive force in the cgs electromagnetic system. The gilbert is a magnetomotive force of $10/4\pi$ ampere-turns. See *magnetomotive force*.

**Hysteresis Loop, Intrinsic.* A hysteresis loop obtained with a ferromagnetic material by plotting (usually to rectangular coordinates) corresponding dc values of intrinsic induction B_i for ordinates and magnetizing force H for abscissas.

**Hysteresis Loop, Normal.* A closed curve obtained with a ferromagnetic material by plotting (usually to rectangular coordinates) corresponding dc values of magnetic induction B for ordinates and magnetizing force H for abscissas when the material is passing through a complete cycle between equal definite limits of either magnetizing force $\pm H_m$ or magnetic induction $\pm B_m$. In general, the normal hysteresis loop has mirror symmetry with respect to the origin of the B and H axes, but this may not be true for special materials.

**Hysteresis-Loop Loss* W_h. The energy expended in a single slow excursion around a normal hysteresis loop is given by the following equation:

$$W_h = \int \frac{HdB}{4\pi} \qquad \text{ergs} \tag{4-39}$$

where the integrated area enclosed by the loop is measured in gauss-oersteds.

***Hysteresis Loss, Normal,* P_h.

1. The power expended in a ferromagnetic material, as a result of hysteresis, when the material is subjected to an *SCM* excitation.

2. The energy loss/cycle in a magnetic material as a result of magnetic hysteresis when the induction is cyclic (but not necessarily periodic).

Hysteresis, Magnetic The property of a ferromagnetic material exhibited by the lack of correspondence between the changes in induction resulting from increasing magnetizing force from decreasing magnetizing force.

Induction B. See *magnetic induction (flux density)*.

**Induction, Intrinsic,* B_i. The vector difference between the magnetic induction in a magnetic material and the magnetic induction that would exist in a vacuum under the influence of the same magnetizing force. This is expressed by the equation

$$B_i = B - \Gamma_m H \tag{4-40}$$

NOTE: In the cgs-em system, $B_i/4\pi$ is often called *magnetic polarization*.

Induction Maximum

*__1.__ B_m—the maximum value of B in a hysteresis loop. The tip of this loop has the magnetostatic coordinates H_m, B_m, which exist simultaneously.

**__2.__ B_{max}—the maximum value of induction, in a flux-current loop.

NOTE: In a flux-current loop, the magnetodynamic values B_{max} and H_{max} do not exist simultaneously; B_{max} occurs later than H_{max}.

**Induction, Normal, B.* The maximum induction in a magnetic material that is in a symmetrically cyclically magnetized condition.

NOTE: Normal induction is a magnetostatic parameter usually measured by hallistic methods.

**Induction, Remanent, B_d.* The magnetic induction that remains in a magnetic circuit after the removal of an applied magnetomotive force.

NOTE: If there are no air gaps or other inhomogeneities in the magnetic circuit, the remanent induction B_r will equal the residual induction B_r; if air gaps or other inhomogeneities are present, B_d will be less than B_r.

**Induction, Residual, B_r.* The magnetic induction corresponding to zero magnetizing force in a magnetic material that is in a symmetrically cyclically magnetized condition.

**Induction, Saturation, B_r.* The maximum intrinsic induction possible in a material.

**Induction Curve, Intrinsic (Ferric).* A curve of a previously demagnetized specimen depicting the relation between intrinsic induction and corresponding ascending values of magnetizing force. This curve starts at the origin of the B_i and H axes.

**Induction Curve, Normal.* A curve of a previously demagnetized specimen depicting the relation between normal induction and corresponding ascending values of magnetizing force. This curve starts at the origin of the B and H axes.

Isotropic Material. Material in which the magnetic properties are the same for all directions.

Magnetic Circuit. A region at whose surface the magnetic induction is tangential.

NOTE: A practical magnetic circuit is the region containing the flux of practical interest, such as the core of a transformer. It may consist of ferromagnetic material with or without air gaps or other feebly magnetic materials such as porcelain and brass.

Magnetic Constant (Permeability of Space) Γ_m. The dimensional scalar factor that relates the mechanical force between two currents to their intensities and geometrical configurations. That is,

$$dF = \Gamma_m I_1 I_2 dl_1 \times \frac{dl_2 \times r_1}{nr^2} \tag{4-41}$$

where Γ_m = magnetic constant when the element of force dF of a current element $I_1\ dl_1$ on another current element $I_2\ dl_2$ is at a distance r

r_1 = unit vector in the direction from dl_1 to dl_2

n = dimensionless factor, the symbol n is unity in unrationalized systems and 4π in rationalized systems

NOTE 1: The numerical values of Γ_m depend on the system of units employed. In the cgs-em system, $\Gamma_m = 1$; in the rationalized mksa system, $\Gamma_m = 4\pi \times 10^{-7}$ h/m.

NOTE 2: The magnetic constant expresses the ratio of magnetic induction to the corresponding magnetizing force at any point in a vacuum and therefore is sometimes called the permeability of space μ_r.

NOTE 3: The magnetic constant times the relative permeability is equal to the absolute permeability:

$$\mu_{abs} = \Gamma_m \mu_r \tag{4-42}$$

Magnetic Field Strength H. See *magnetizing force.*

Magnetic Flux ϕ. The product of the magnetic induction B and the area of a surface (or cross section) A when the magnetic induction B is uniformly distributed and normal to the plane of the surface.

$$\phi = BA \tag{4-43}$$

where ϕ = magnetic flux
B = magnetic induction
A = area of the surface

NOTE 1: If the magnetic induction is not uniformly distributed over the surface, the flux ϕ is the surface integral of the normal component of B over the area:

$$\phi = \iint_s B\, dA \tag{4-44}$$

NOTE 2: Magnetic flux is scalar and has no direction.

Magnetic Flux Density B. See *magnetic induction (flux density).*

Magnetic Induction (Flux Density) B. That magnetic vector quantity which at any point in a magnetic field is measured either by the mechanical force experienced by an element of electric current at the point, or by the electromotive force induced in an elementary loop during any change in flux linkages with the loop at the point.

NOTE 1: If the magnetic induction B is uniformly distributed and normal to a surface or cross section, then the magnetic induction is

$$B = \phi/A \tag{4-45}$$

where B = magnetic induction
ϕ = total flux
A = area

NOTE 2: B_{in} is the instantaneous value of the magnetic induction and B_m is the maximum value of the magnetic induction.

Magnetizing Force (Magnetic Field Strength) H. That magnetic vector quantity at a point in a magnetic field which measures the ability of electric currents or magnetized bodies to produce magnetic induction at the given point.

NOTE 1: The magnetizing force H may be calculated from the current and the geometry of certain magnetizing circuits. For example, in the center of a uniformly wound long solenoid,

$$H = C(NI/l) \tag{4-46}$$

where H = magnetizing force
C = constant whose value depends on the system of units
N = number of turns
I = current
l = axial length of the coil

If I is expressed in amperes and l is expressed in centimeters, then $C = 4\pi/10$ in order to obtain H in the cgs = em unit, the oersted. If I is expressed in amperes and l is expressed in meters, then $C = 1$ in order to obtain H in the mksa unit, ampere-turn per meter.

NOTE 2: The magnetizing force H at a point in air may be calculated from the measured value of induction at the point by dividing this value by the magnetic constant Γ_m.

***Magnetizing Force, AC.* Three different values of dynamic magnetizing force parameters are in common use:

1. H_L—an assumed peak value computed in terms of peak magnetizing current (considered to be sinusoidal).
2. H_x—an assumed peak value computed in terms of measured rms exciting current (considered to be sinusoidal).
3. H_p—computed in terms of a measured peak value of exciting current, and thus equal to the value H'_{max}.

***Magnetodynamic.* The magnetic condition when the values of magnetizing force and induction vary, usually periodically and repetitively, between two extreme limits.

Magnetomotive Force F. The line integral of the magnetizing force around any flux loop in space.

$$F = \oint H\,dl \tag{4-47}$$

where F = magnetomotive force
H = magnetizing force
dl = unit length along the loop

NOTE: The magnetomotive force is proportional to the net current linked with any closed loop of flux or closed path

$$F = CNI \tag{4-48}$$

where F = magnetomotive force
N = number of turns linked with the loop
I = current in amperes
C = constant whose value depends on the system of units. In the cgs system, $C = 4\pi/10$. In the mksa system, $C = 1$

**Magnetostatic.* The magnetic condition when the values of magnetizing force and induction are considered to remain invariant with time during the period of measurement. This is often referred to as a dc (direct-current) condition.

Magnetostriction. Changes in dimensions of a body resulting from magnetization.

Maxwell. The unit of magnetic flux in the cgs electromagnetic system. One maxwell equals 10^{-8} weber. See *magnetic flux.*

NOTE:

$$e = -N\frac{d\phi}{dt} \times 10^{-8} \tag{4-49}$$

where e = induced instantaneous emf volts
$d\phi/dt$ = time rate of change of flux, maxwells per second
N = number of turns surrounding the flux, assuming each turn is linked with all the flux

Oersted. The unit of magnetizing force (magnetic field strength) in the cgs electromagnetic system. One oersted equals a magnetomotive force of 1 gilbert/cm of flux path. One oersted equals $100/4\pi$ or 79.58 ampere-turns per meter. See *magnetizing force (magnetic field strength).*

Paramagnetic Material. A material having a relative permeability which is slightly greater than unity, and which is practically independent of the magnetizing force.

***Permeability, AC.* A generic term used to express various dynamic relationships between magnetic induction B and magnetizing force H for magnetic material subjected to a cyclic excitation by alternating or pulsating current. The values of ac permeability obtained for a given material depend fundamentally on the excursion limits of dynamic excitation and induction, the method and conditions of measurement, and also on such factors as resistivity, thickness of laminations, frequency of excitation, etc.

NOTE: The numerical value for any permeability is meaningless unless the corresponding B or H excitation level is specified. For incremental permeabilities, not only the corresponding dc B or H excitation level must be specified but also the dynamic excursion limits of dynamic excitation range (ΔB or ΔH).

AC permeabilities in common use for magnetic testing are

1. ***Impedance (rms) permeability* μ_z. The ratio of the measured peak value of magnetic induction to the value of the apparent magnetizing force H_z calculated from the measured rms value of the exciting current, for a material in the *SCM* condition.

NOTE: The value of the current used to compute H_z is obtained by multiplying the measured value of rms exciting current by 1.414. This assumes that the total exciting current is magnetizing current and is sinusoidal.

2. ***Inductance permeability* μ_L. For a material in an *SCM* condition, the permeability is evaluated from the measured inductive component of the electric circuit representing the magnetic specimen. This circuit is assumed to be composed of paralleled linear inductive and resistive elements ωL_1 and R_1.

3. ***Peak permeability* μ_p. The ratio of the measured peak value of magnetic induction to the peak value of the magnetizing force H_p, calculated from the measured peak value of the exciting current, for a material in the *SCM* condition.

Other ac permeabilities are:

4. *Ideal permeability* μ_a. The ratio of the magnetic induction to the corresponding magnetizing force after the material has been simultaneously subjected to a value of ac magnetizing force approaching saturation (of approximate sine waveform) superimposed on a given dc magnetizing force, and the ac magnetizing force has thereafter been gradually reduced to zero. The resulting ideal permeability is thus a function of the dc magnetizing force used.

NOTE: Ideal permeability, sometimes called anhysteretic permeability, is principally significant to feebly magnetic material and to the Rayleigh range of soft magnetic material.

5. ***Impedance, permeability, incremental*, $\mu_{\Delta z}$. Impedance permeability μ_z obtained when an ac excitation is superimposed on a dc excitation, *CM* condition.

6. ***Inductance permeability, incremental*, $\mu_{\Delta L}$. Inductance permeability μ_L obtained when an ac excitation is superimposed on a dc excitation, *CM* condition.

7. ***Initial dynamic permeability* μ_{0d}. The limiting value of inductance permeability μ_L reached in a ferromagnetic core when, under *SCM* excitation, the magnetizing current has been progressively and gradually reduced from a comparatively high value to zero value.

NOTE: This same value, μ_{0d}, is also equal to the initial values of both impedance permeability μ_x and peak permeability μ_p.

8. ***Instantaneous permeability (coincident with* B_{max}*)* μ_t. With *SCM* excitation, the ratio of the maximum induction B_{max} to the instantaneous magnetizing force H_t, which is the value of apparent magnetizing force H' determined at the instant when B reaches a maximum.

9. ***Peak permeability, incremental*, $\mu_{\Delta p}$. Peak permeability μ_p obtained when an ac excitation is superimposed on dc excitation, *CM* condition.

**Permeability, DC.* Permeability is a general term used to express relationships between magnetic induction B and magnetizing force H under various conditions of magnetic excitation. These relationships are either (1) absolute permeability, which in general is the quotient of a change in magnetic induction divided by the corresponding change in magnetizing force, or (2) relative permeability, which is the ratio of the absolute permeability to the magnetic constant Γ_m.

NOTE 1: The magnetic constant Γ_m is a scalar quantity differing in value and uniquely determined by each electromagnetic system of units. In the unrationalized cgs system, Γ_m is 1 gauss/oersted and in the mksa rationalized system $\Gamma_m = 4\pi \times 10^{-7}$ H/m.

NOTE 2: Relative permeability is a pure number which is the same in all unit systems. The value and dimension of absolute permeability depend on the system of units employed.

NOTE 3: For any ferromagnetic material, permeability is a function of the degree of magnetization. However, initial permeability μ_0 and maximum permeability μ_m are unique values for a given specimen under specified conditions.

NOTE 4: Except for initial permeability μ_0, a numerical value for any of the dc permeabilities is meaningless unless the corresponding B or H excitation level is specified.

NOTE 5: For the incremental permeabilities μ_Δ and $\mu_{\Delta i}$, a numerical value is meaningless unless both the corresponding values of mean excitation level ($\bar{B}$ or $\bar{H}$) and the excursion range (ΔB or ΔH) are specified.

The following dc permeabilities are frequently used in magnetostatic measurements primarily concerned with the testing of materials destined for use with permanent or dc excited magnets:

1. **Absolute permeability* μ_{abs}. The sum of the magnetic constant and the intrinsic permeability. It is also equal to the product of the magnetic constant and the relative permeability.

$$\mu_{\text{abs}} = \Gamma_m + \mu_i = \Gamma_m \mu_r \tag{4-50}$$

2. **Differential permeability* μ_d. The absolute value of the slope of the hysteresis loop at any point, or the slope of the normal magnetizing curve at any point.

3. **Effective circuit permeability* μ_{eff}. When a magnetic circuit consists of two or more components, each individually homogeneous throughout but having different permeability values, the effective (overall) permeability of the circuit is that value computed in terms of the total magnetomotive force, the total resulting flux, and the geometry of the circuit.

NOTE: For a symmetrical series circuit in which each component has the same cross-sectional area, reluctance values add directly, giving

$$\mu_{\text{eff}} = \frac{l_1 + l_2 + l_3 + \cdots}{l_1/\mu_1 + l_2/\mu_2 + l_3/\mu_3 + \cdots} \tag{4-51}$$

For a symmetrical parallel circuit in which each component has the same flux path length, permeance values add directly, giving

$$\mu_{\text{eff}} = \frac{\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 + \cdots}{A_1 + A_2 + A_3 + \cdots} \tag{4-52}$$

4. **Incremental intrinsic permeability* $\mu_{\Delta i}$. The ratio of the change in intrinsic induction to the corresponding change in magnetizing force when the mean induction differs from zero.

5. **Incremental permeability* μ_Δ. The ratio of a change in magnetic induction to the corresponding change in magnetizing force when the mean induction differs from zero. It equals the slope of a straight line joining the excursion limits of an incremental hysteresis loop.

NOTE: When the change in H is reduced to zero, the incremental permeability μ_Δ becomes the reversible permeability μ_{rev}.

6. **Initial permeability* μ_0. The limiting value approached by the normal permeability as the applied magnetizing force H is reduced to zero. The permeability is equal to the slope of the normal induction curve at the origin of linear B and H axes.

7. **Intrinsic permeability* μ_i. The ratio of intrinsic induction to the corresponding magnetizing force.

8. **Maximum permeability* μ_m. The value of normal permeability for a given material where a straight line from the origin of linear B and H axes becomes tangent to the normal induction curve.

9. **Normal permeability* μ *(without subscript)*. The ratio of the normal induction to the corresponding magnetizing force. It is equal to the slope of a straight line joining the extrusion limits of a normal hysteresis loop, or the slope of a straight line joining any point (H_m, B_m) on the normal induction curve to the origin of the linear B and H axes.

10. **Relative permeability* μ_r. The ratio of the absolute permeability of a material to the magnetic constant Γ_m giving a pure numeric parameter.

NOTE: In the cgs-em system of units, the relative permeability is numerically the same as the absolute permeability.

11. *Reversible permeability* μ_{rev}. The limit of the incremental permeability as the change in magnetizing force approaches zero.

12. *Space permeability* μ_o. The permeability of space (vacuum), identical with the magnetic constant Γ_m.

***Reactive Power (Quadrature Power)* P_q. The product of the rms current in an electric circuit, the rms voltage across the circuit, and the sine of the angular phase difference between the current and the voltage.

$$P_q = EI \sin \theta \tag{4-53}$$

where P_q = reactive power, vars
E = voltage, volts
I = current, amperes
θ = angular phase by which E leads I

NOTE: The reactive power supplied to a magnetic core having an *SCM* excitation is the product of the magnetizing current and the voltage induced in the exciting winding.

**Remanence* B_{dm}. The maximum value of the remanent induction for a given geometry of the magnetic circuit.

NOTE: If there are no air gaps or other inhomogeneities in the magnetic circuit, the remanence B_{dm} is equal to the retentivity B_{rs}; if air gaps or other inhomogeneities are present, B_{dm} will be less than B_{rs}.

**Retentivity* B_{rs}. That property of a magnetic material which is measured by its maximum value of the residual induction.

NOTE: Retentivity is usually associated with saturation induction.

Symmetrically Cyclically Magnetized Condition, SCM. A magnetic material is in an *SCM* condition when, under the influence of a magnetizing force that varies cyclically between two equal positive and negative limits, its successive hysteresis loops or flux-current loops are both identical and symmetrical with respect to the origin of the axes.

Tesla. The unit of magnetic induction in the mksa (Giorgi) system. The tesla is equal to 1 Wb/m^2 or 10^4 gausses.

Var. The unit of reactive (quadrature) power in the mksa (Giorgi) and the practical systems.

Volt-Ampere. The unit of apparent power in the mksa (Giorgi) and the practical systems.

Watt. The unit of active power in the mksa (Giorgi) and the practical systems. One watt is a power of 1 J/s.

Weber. The unit of magnetic flux in the mksa and in the practical system. The weber is the magnetic flux whose decrease to zero when linked with a single turn induces in the turn a voltage whose time integral is 1 v/s. One weber equals 10^8 maxwells. See *magnetic flux*.

4.2.2 Magnetic Properties and Their Application

The relative importance of the various magnetic properties of a magnetic material varies from one application to another. In general, properties of interest may include normal induction, hysteresis, dc permeability, ac permeability, core loss, and exciting power. It should be noted that there are various means of expressing ac permeability. The choice depends primarily on the ultimate use.

Techniques for the magnetic testing of many magnetic materials are described in the ASTM standards. The magnetic and electric circuits employed in magnetic testing of a specimen are as free as possible from any unfavorable design factors which would prevent the measured magnetic data from being representative of the inherent magnetic properties of the specimen. The flux "direction" in the specimen is normally specified, since most magnetic materials are magnetically anisotropic. In most ac magnetic tests, the waveform of the flux is required to be sinusoidal.

As a result of the existence of unfavorable conditions, such as those listed and described below, the performance of a magnetic material in a magnetic device can be greatly deteriorated from that which would be expected from magnetic testing of the material. Allowances for these conditions, if present, must be made during the design of the device if the performance of the device is to be correctly predicted.

Leakage. A principal difficulty in the design of many magnetic circuits is due to the lack of a practicable material which will act as an insulator with respect to magnetic flux. This results in magnetic flux seldom being completely confined to the desired magnetic circuit. Estimates of leakage flux for a particular design may be made based on experience and/or experimentation.

Flux Direction. Some magnetic materials have a very pronounced directionality in their magnetic properties. Failure to utilize these materials in their preferred directions results in impaired magnetic properties.

Fabrication. Stresses introduced into magnetic materials by the various fabricating techniques often adversely affect the magnetic properties of the materials. This occurs particularly in materials having high permeability. Stresses may be eliminated by a suitable stress-relief anneal after fabrication of the material to final shape.

Joints. Joints in an electromagnetic core may cause a large increase in total excitation requirements. In some cores operated on ac, core loss may also be increased.

Waveform. When a sinusoidal voltage is applied to an electromagnetic core, the resulting magnetic flux is not necessarily sinusoidal in waveform, especially at high inductions. Any harmonics in the flux waveform cause increases in core loss and required excitation power.

Flux Distribution. If the maximum and minimum lengths of the magnetic path in an electromagnetic core differ too much, the flux density may be appreciably greater at the inside of the core structure than at the outside. For cores operated on ac, this can cause the waveform of the flux at the extremes of the core structure to be distorted even when the total flux waveform is sinusoidal.

4.2.3 Types of Magnetism

Any substance may be classified into one of the following categories according to the type of magnetic behavior it exhibits:

1. Diamagnetic
2. Paramagnetic
3. Antiferromagnetic
4. Ferromagnetic
5. Ferrimagnetic

Substances that fall into the first three categories are so weakly magnetic that they are commonly thought of as *nonmagnetic*. In contrast, ferromagnetic and ferrimagnetic substances are strongly magnetic and are thereby of interest as *magnetic materials*. The magnetic behavior of any ferromagnetic or ferrimagnetic material is a result of its spontaneously magnetized magnetic domain structure and is characterized by a nonlinear normal induction curve, hysteresis, and saturation.

The pure elements which are ferromagnetic are iron, nickel, cobalt, and some of the rare earths. Ferromagnetic materials of value to industry for their magnetic properties are almost invariably alloys of the metallic ferromagnetic elements with one another and/or with other elements.

Ferrimagnetism occurs mainly in the ferrites, which are chemical compounds having ferric oxide (Fe_2O_3) as a component. In recent years, some of the magnetic ferrites have become very important in certain magnetic applications. The magnetic ferrites saturate magnetically at lower inductions than do the great majority of metallic ferromagnetic materials. However, the electrical resistivities of ferrites are at least several orders of magnitude greater than those of metals.

Commercial Magnetic Materials. *Commercial magnetic materials* are generally divided into two main groups, each composed of ferromagnetic and ferrimagnetic substances:

1. Magnetically "soft" materials
2. Magnetically "hard" materials

The distinguishing characteristic of "soft" magnetic materials is high permeability. These materials are employed as core materials in the magnetic circuits of electromagnetic equipment. "Hard" magnetic materials are characterized by a high maximum magnetic energy product BH_{max}. These materials are employed as permanent magnets to provide a constant magnetic field when it is inconvenient or uneconomical to produce the field by electromagnetic means.

4.2.4 "Soft" Magnetic Materials

A wide variety of "soft" magnetic materials have been developed to meet the many different requirements imposed on magnetic cores for modern electrical apparatus and electronic devices. The various soft magnetic materials will be considered under three classifications:

1. Materials for solid cores.
2. Materials for laminated cores.
3. Materials for special purposes.

4.2.5 Materials for Solid Cores

These materials are used in dc applications such as yokes of dc dynamos, rotors of synchronous dynamos, and cores of dc electromagnets and relays. Proper annealing of these materials improves their magnetic properties. The principal magnetic requirements for the solid-core materials are high saturation, high permeability at relatively high inductions, and at times, low coercive force.

Wrought iron is a ferrous material, aggregated from a solidifying mass of pasty particles of highly refined metallic iron, into which is incorporated, without subsequent fusion, a minutely and uniformly distributed quantity of slag. The better types of wrought iron are known as *Norway iron* and *Swedish iron* and are widely used in relays after being annealed to reduce coercive force and to minimize magnetic aging.

Cast irons are irons which contain carbon in excess of the amount which can be retained in solid solution in austenite at the eutectic temperature. The minimum carbon content is about 2%, while the practical maximum carbon content is about 4.5%. Cast iron was used in the yokes of dc dynamos in the early days of such machines.

Gray cast iron is a cast iron in which graphite is present in the form of flakes. It has very poor magnetic properties, inferior mechanical properties, and practically no ductility. It does lend itself well to the casting of complex shapes and is readily machinable.

Malleable cast iron is a cast iron in which the graphite is present as temper carbon nodules. It is magnetically better than gray cast iron.

Ductile (nodular) cast iron is a cast iron with the graphite essentially spheroidal in shape. It is magnetically better than gray cast iron. Ductile cast iron has the good castability and machinability of gray cast iron together with much greater strength, ductility, and shock resistance.

4.2.6 Carbon Steels

Carbon steels may contain from less than 0.1% carbon to more than 1% carbon. The magnetic properties of a carbon steel are greatly influenced by the carbon content and the disposition of the carbon. Low-carbon steels (less than 0.2% carbon) have magnetic properties which are similar to those of wrought iron and far superior to those of any of the cast irons.

Wrought carbon steels are widely used as solid-core materials. The low-carbon types are preferred in most applications.

Cast carbon steels replaced cast iron many years ago as the material used in the yokes of dc machines, but have since largely been supplanted in this application by wrought (hot-rolled) carbon-steel plates of welding quality.

4.2.7 Materials for Laminated Cores

The materials most widely employed in wound or stacked cores in electromagnetic devices operated at the commercial power frequencies (50 and 60 Hz) are the electrical steels and the specially processed carbon steels designated as *magnetic lamination steels*. The principal magnetic requirements for these materials are low core loss, high permeability, and high saturation. ASTM publishes standard specifications for these materials. On a tonnage basis, production of these materials far exceeds that of any other magnetic material.

Electrical steels are flat-rolled low-carbon silicon-iron alloys. Since applications for electrical steels lie mainly in energy-loss-limited equipment, the core losses of electrical steels are normally guaranteed by the producers. The general category of electrical steels may be divided into classifications of (1) nonoriented materials and (2) grain-oriented materials.

Electrical steels are usually graded by high-induction core loss. Both ASTM and AISI have established and published designation systems for electrical steels based on core loss.

The ASTM core loss type designation consists of six or seven characters. The first two characters are 100 times the nominal thickness of the material in millimeters. The third character is a code letter which designates the class of the material and specifies the sampling and testing practices. The last three or four characters are 100 times the maximum permissible core loss in watts per pound at a specified test frequency and induction.

The AISI designation system has been discontinued but is still widely used. The AISI type designation for a grade consisted of the letter M followed by a number. The letter M stood for magnetic material, and the number was approximately equal to 10 times the maximum permissible core loss in watts per pound for 0.014-in material at 15 kG, 60 Hz in 1947.

Nonoriented electrical steels have approximately the same magnetic properties in all directions in the plane of the material (see Figs. 4-8 and 4-9). The common application is in punched laminations for large

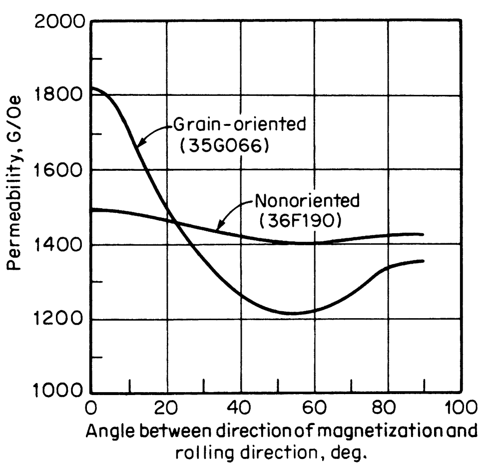


FIGURE 4-8 Effect of direction of magnetization on normal permeability at 10 Oe of fully processed electrical steels.

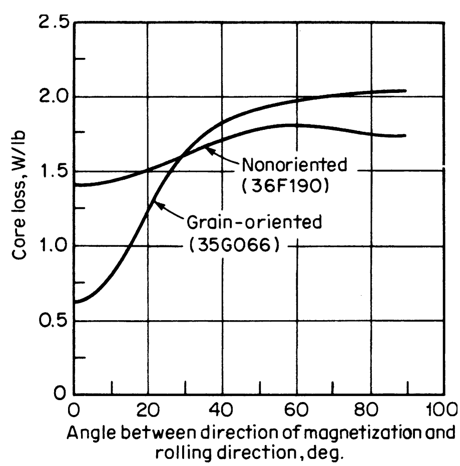


FIGURE 4-9 Effect of direction of magnetization on core loss at 15 kG, 60 Hz or fully processed electrical steel.

and small rotating machines and for small transformers. Today, nonoriented materials are always cold-rolled to final thickness. Hot rolling to final thickness is no longer practiced. Nonoriented materials are available in both fully processed and semiprocessed conditions.

Fully processed nonoriented materials have their magnetic properties completely developed by the producer. Stresses introduced into these materials during fabrication of magnetic cores must be relieved by annealing to achieve optimal magnetic properties in the cores. In many applications, however, the degradation of the magnetic properties during fabrication is slight and/or can be tolerated, and the stress-relief anneal is omitted. Fully processed nonoriented materials contain up to about 3.5% silicon. Additionally, a small amount (about 0.5%) of aluminum is usually present. The common thicknesses are 0.014, 0.0185, and 0.025 in.

Semiprocessed nonoriented materials do not have their inherent magnetic properties completely developed by the producer and must be annealed properly to achieve both decarburization and grain growth. These materials are used primarily in high-volume production of small laminations and cores which would require stress-relief annealing if made from fully processed material. Semiprocessed nonoriented materials contain up to about 3% silicon. Additionally, a small amount (about 0.5%) of aluminum is usually present. The carbon content may be as high as 0.05% but should be reduced to 0.005% or less by the required anneal. The common thicknesses of semiprocessed nonoriented materials are 0.0185 and 0.025 in.

Grain-oriented electrical steels have a pronounced directionality in their magnetic properties (Figs. 4-8 and 4-9). This directionality is a result of the "cube-on-edge" crystal structure achieved by proper composition and processing. Grain-oriented materials are employed most effectively in magnetic cores in which the flux path lies entirely or predominantly in the rolling direction of the material. The common application is in cores of power and distribution transformers for electric utilities.

Grain-oriented materials are produced in a fully processed condition, either unflattened or thermally flattened, in thicknesses of 0.0090, 0.0106, 0.0118, and 0.0138 in. Unflattened material has appreciable coil set or curvature. It is used principally in making spirally wound or formed cores. These cores must be stress-relief annealed to relieve fabrication stresses. Thermally flattened material is employed principally in making sheared or stamped laminations. Annealing of the laminations to remove both residual stresses from the thermal-flattening and fabrication stresses is usually recommended. However, special thermally flattened materials are available which do not require annealing when used in the form of wide flat laminations.

Two types of grain-oriented electrical steels are currently being produced commercially. The regular type, which was introduced many years ago, contains about 3.15% silicon and has grains about 3 mm in diameter. The high-permeability type, which was introduced more recently, contains about 2.9% silicon and has grains about 8 mm in diameter. In comparison with the regular type, the high-permeability type has better core loss and permeability at high inductions.

Some characteristics and applications for electrical steels are shown in Table 4-4.

Surface insulation of the surfaces of electrical steels is needed to limit the interlaminar core losses of magnetic cores made of electrical steels. Numerous surface insulations have been developed to meet the requirements of various applications. The various types of surface insulations have been classified by AISI.

Annealing of laminations or cores made from electrical steels is performed to accomplish either stress relief in fully processed material or decarburization and grain growth in semiprocessed material. Both batch-type annealing furnaces and continuous annealing furnaces are employed. The former is best suited for low-volume or varied production, while the latter is best suited for high-volume production.

Stress-relief annealing is performed at a soak temperature in the range from 730 to 845°C. The soak time need be no longer than that required for the charge to reach soak temperature. The heating and cooling rates must be slow enough so that excessive thermal gradients in the material are avoided. The annealing atmosphere and other annealing conditions must be such that chemical contamination of the material is avoided.

Annealing for decarburization and grain growth is performed at a soak temperature in the range from 760 to 870°C. Atmospheres of hydrogen or partially combusted natural gas and containing water vapor are often used. The soak time required for decarburization depends not only on the

TABLE 4-4 Some Characteristics and Typical Applications for Specific Types of Electrical Steels

ASTM type	Some characteristics	Typical applications
	Oriented types	
23G048 through 35G066 or 27H076 through 35H094 or 27P066 through 35P076	Highly directional magnetic properties due to grain orientation. Very low core loss and high permeability in rolling direction.	Highest-efficiency power and distribution transformers with lower weight per kVA. Large generators and power transformers.
	Nonoriented types	
36F145 and 47F168	Lowest core loss, conventional grades. Excellent permeability at low inductions.	Small power transformers and rotating machines of high efficiency.
36F158 through 64F225 or 47S178 and 64S194	Low core loss, good permeability at low and intermediate inductions.	High-reactance cores, generators, stators of high-efficiency rotating equipment.
36F190 through 64F270 or 47S188 through 64S260	Good core loss, good permeabilty at all inductions, and low exciting current. Good stamping properties.	Small generators, high-efficiency, continuous duty rotating ac and dc machines.
47F290 through 64F600 or 47S250 through 64S350	Ductile, good stamping properties, good permeabilty at high inductions.	Small motors, ballasts, and relays.

temperature and atmosphere but also on the dimensions of the laminations or cores being annealed. If the dimensions are large, long soak times may be required.

Magnetic lamination steels are cold-rolled low-carbon steels intended for magnetic applications, primarily at power frequencies. The magnetic properties of magnetic lamination steels are not normally guaranteed and are generally inferior to those of electric steels. However, magnetic lamination steels are frequently used as core materials in small electrical devices, especially when the cost of the core material is a more important consideration than the magnetic performance.

Usually, but not always, stamped laminations or assembled core structures made from magnetic lamination steels are given a decarburizing anneal to enhance the magnetic properties. Optimal magnetic properties are obtained when the carbon content is reduced to 0.005% or less from its initial value, which may approach 0.1%. The soak temperature of the anneal is in the range from 730 to 790°C. The atmosphere most often used at the present time is partially combusted natural gas with a suitable dew point. Soak time depends to a considerable degree on the dimensions of the laminations or core structures being annealed.

Three types of magnetic lamination steels are produced. Type 1 is usually made to a controlled chemical composition and is furnished in the full-hard or annealed condition without guaranteed magnetic properties. Type 2 is made to a controlled chemical composition, given special processing, and furnished in the annealed condition without guaranteed magnetic properties. After a suitable anneal, the magnetic properties of Type 2 are superior to those of Type 1. Type 2S is similar to Type 2, but the core loss is guaranteed.

4.2.8 Materials for Special Purposes

For certain applications of soft or nonretentive materials, special alloys and other materials have been developed, which, after proper fabrication and heat treatment, have superior properties in certain ranges of magnetization. Several of these alloys and materials will be described.

Nickel-Iron Alloys. Nickel alloyed with iron in various proportions produces a series of alloys with a wide range of magnetic properties. With 30% nickel, the alloy is practically nonmagnetic and has a resistivity of 86 $\mu\Omega$/cm. With 78% nickel, the alloy, properly heat-treated, has very high permeability. These effects are shown in Figs. 4-10 and 4-11. Many variations of this series have been

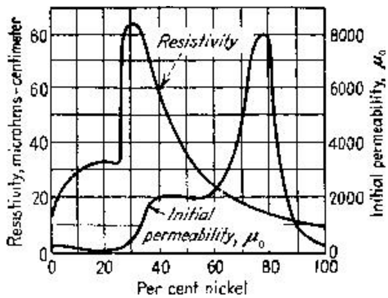


FIGURE 4-10 Electrical resistivity and initial permeability of iron-nickel alloys with various nickel contents.

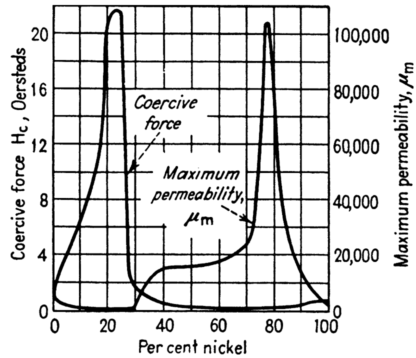


FIGURE 4-11 Maximum permeability and coercive force of iron-nickel alloys with various nickel contents.

developed for special purposes. Table 4-5 lists some of the more important commercial types of nickel-iron alloys, with their approximate properties. These alloys are all very sensitive to heat treatment, so their properties are largely influenced thereby.

Permalloy. This is a term applied to a number of nickel-iron alloys developed by the Bell Laboratories, each specified by a prefix number indicating the nickel content. The term is usually associated with the 78.5% nickel-iron alloys, the important properties of which are high permeability and low hysteresis loss in relatively low magnetizing fields. These properties are obtained by a unique heat treatment consisting of a high-temperature anneal, preferably in hydrogen, with slow cooling followed by rapid cooling from about 625°C. The alloy is very sensitive to mechanical strain, so it is desirable to heat-treat the alloy in its final form. The addition of 3.8% chromium or molybdenum increases the resistivity from 16 to 65 and 55 $\mu\Omega \cdot$ cm, respectively, without seriously impairing the magnetic quality. In fact, low-density permeabilities are better with these additions. These alloys have found their principal application as a material for the continuous loading of submarine cables and in loading coils for landlines.

By special long-time high temperature treatments, maximum permeability values greater than 1 million have been obtained. The double treatment required by the 78% Permalloy is most effective when the strip is thin, say, under 10 mils. For greater thicknesses, the quick cooling from 625°C is not uniform throughout the section, and loss of quality results.

TABLE 4-5 Special-Purpose Materials

Name	Approximate composition, %	Saturation, G	Maximum permeability	Coercivity (from saturation), Oe	Initial permeability	Resistivity, $\mu\Omega \cdot$ cm
78 Permalloy	78.5 Ni	10,500	70,000	–0.05	8,000	16
MoPermalloy	79 Ni, 4.0 Mo	8,000	90,000	–0.05	20,000	55
Supermalloy	79 Ni, 5 Mo	7,900	900,000	–0.002	100,000	60
48% nickel-iron	48 Ni	16,000	60,000	–0.06	5,000	45
Monimax	47 Ni, 3 Mo	14,500	35,000	–0.10	2,000	80
Sinimax	43 Ni, 3 Si	11,000	35,000	–0.10	3,000	85
Mumetal	77 Ni, 5 Cu, 2 Cr	6,500	85,000	–0.05	20,000	60
Deltamax	50 Ni	15,500	85,000	–0.10		45

A 48% nickel-iron was developed for applications requiring a moderately high-permeability alloy with higher saturation density than 78 Permalloy. The same general composition is marketed under many names, such as Hyperm 50, Hipernik, Audiolloy, Allegheny Electric Metal, 4750, and Carpenter 49 alloy. Annealing is recommended after all mechanical operations are completed. These alloys have found extensive use in radio, radar, instrument, and magnetic-amplifier components.

Deltamax. By the use of special techniques of cold reduction and annealing, the 48% nickel-iron alloy develops directional properties resulting in high permeability and a square hysteresis loop in the rolling direction. A similar product is sold under the name of Orthonic. For optimal properties, these materials are rapidly cooled after a 2-h anneal in pure hydrogen at 1100°C. They are generally used in wound cores of thin tape for applications such as pulse transformers and magnetic amplifiers.

Iron-Nickel-Copper-Chromium. The addition of copper and chromium to high-nickel-iron alloys has the effect of raising the permeability at low flux density. Alloys of this type are marketed under the names of Mumetal, 1040 alloy, and Hymu 80. For optimal properties, they are annealed after cutting and forming for 4 h at 1100°C in pure hydrogen and cooled slowly. Important applications are as magnetic shielding for instruments and electronic equipment and as cores in magnetic amplifiers.

Constant-Permeability Alloys. *Constant-permeability alloys* having a moderate permeability which is quite constant over a considerable range of flux densities are desirable for use in circuits in which waveform distortion must be kept at a minimum. Isoperm and Conpernik are two alloys of this type. They are nickel-iron alloys containing 40% to 55% nickel which have been severely cold-worked. Perminvar is the name given to a series of cobalt-nickel-iron alloys (e.g., 50% nickel, 25% cobalt, 25% iron) which also exhibit this characteristic of constant permeability over a low (~800 G) density range. When magnetized to higher flux densities, they give a double loop constricted at the origin so as to give no measurable remanence or coercive force. The characteristics of the alloys in this group vary greatly with the chemical content and the heat treatment. A sample containing approximately 45 Ni, 25 Co, and 30 Fe, baked for 24 h at 425°C and slowly cooled, had hysteresis losses as follows: At 100 G, 214×10^{-4} erg/(cm^3)(cycle); at 1003 G, 15.27 ergs; at 1604 G, 163 ergs; at 4950 G, 1736 ergs; and at 13,810 G, 4430 ergs. Over the range of flux densities in which the permeability is constant (from 0 to 600 G), the hysteresis loss is very small, or on the order of the foregoing figure for 100 G. The resistivity of the sample was 19.63 $\mu\Omega \cdot$ cm.

Monel. *Monel metal* is an alloy of 67% nickel, 28% copper, and 5% other metals. It is slightly magnetic below 95°C.

Iron-Cobalt Alloys. The addition of cobalt to iron has the effect of raising the saturation intensity of iron up to about 36% cobalt (Fe_2Co). This alloy is useful for pole pieces of electromagnets and for any application where high magnetic intensity is desired. It is workable hot but quite brittle cold. *Hyperco* contains approximately ⅓ Co, ⅔ Fe, plus 1% to 2% "added element." Total core loss is about 2.5 W/lb at 15 kG and 0.010 in thick. It is available as hot-rolled sheet, cold-rolled strip, plates, and forgings. The 50% cobalt-iron alloy Permendur has a high permeability in fields up to 50 Oe and, with about 2% vanadium added, can be cold-rolled.

Iron-Silicon Aluminum Alloys. Aluminum in small percentages, usually under 0.5%, is a valuable addition to the iron-silicon alloy. Its principal function appears to be as a deoxidizer. Masumoto has investigated soft magnetic alloys containing much higher percentages of aluminum and found several that have high permeabilities and low hysteresis losses. Certain compositions have very low magnetostriction and anisotropy, high initial permeability, and high electrical resistivity. An alloy of 9.6% silicon and 6% aluminum with iron has better low-flux-density properties than the Permalloys. However, poor ductility has limited these alloys to dc applications in cast configurations or in insulated pressed-powder cores for high-frequency uses. These alloys are commonly known as Sendust. The material has been prepared in sheet form by special processes.

Temperature-Sensitive Alloys. Inasmuch as the Curie point of metal may be moved up or down the temperature scale by the addition of other elements, it is possible to select alloys which lose their ferromagnetism at almost any desired temperature up to 1115°C, the change point in cobalt. Iron-based alloys are ordinarily used to obtain the highest possible permeability at points below the Curie temperature. Nickel, manganese, chromium, and silicon are the most effective alloy elements for this purpose, and most alloys made for temperature-control applications, such as instruments, reactors, and transformers, use one or more of these. The Carpenter Temperature Compensator 30 is a nickel-copper-iron alloy which loses its magnetism at 55°C and is used for temperature compensation in meters.

Heusler's Alloys. *Heusler's alloys* are ferromagnetic alloys composed of "nonmagnetic" elements. Copper, manganese, and aluminum are frequently used as the alloying elements. The saturation induction is about one-third that of pure iron.

4.2.9 High-Frequency Materials Applications

Magnetic materials used in reactors, transformers, inductors, and switch-mode devices are selected on the basis of magnetic induction, permeability, and associated material power losses at the design frequency. Control of eddy currents becomes of primary importance to reduce losses and minimize skin effect produced by eddy-current shielding. This is accomplished by the use of high-permeability alloys in the form of wound cores of thin tape, or compressed, insulated powder iron alloy cores, or sintered ferrite cores.

Typically, the thin magnetic strip material is used in applications where operating frequencies range from 400 Hz to 20 kHz. Power conditioning equipment frequently operates at 10 kHz and up, and the magnetic materials used are compressed, powdered iron-alloy cores or sintered ferrite cores. Power losses in magnetic materials are of great concern, especially so when operated at high frequencies.

3% Silicon-Iron Alloys. *3% Silicon-iron alloys* for high-frequency use are available in an insulated 0.001- to 0.006-in-thick strip that exhibits high effective permeability and low losses at relatively high flux densities. This alloy, as well as other rolled-to-strip soft magnetic alloys, is used to make laminated magnetic cores by various methods, including (1) the wound-core approach for winding toroids and C and E cores, (2) stamped or sheared-to-length laminations for laid-up transformers, and (3) stamped laminations of various configurations (rings E, I, F, L, DU, etc.) for assembly into transformer cores. Laminated core materials usually are annealed after all fabricating and stamping operations have been completed in order to develop the desired magnetic properties of the material. Subsequent forming, bending, or machining may impair the magnetic characteristics developed by the anneal.

Amorphous Metal Alloys. *Amorphous metal alloys* are made using a new technology which produces a thin (0.001 to 0.003 in) ribbon from rapidly quenched molten metal. The alloy solidifies before the atoms have a chance to segregate or crystallize, resulting in a glasslike atomic structure material of high electrical resistivity, 125 to 130 $\mu\Omega \cdot$ cm. A range of magnetic properties may be developed in these materials by using different alloying elements. Amorphous metal alloys may be used in the same high-frequency applications as the cast, rolled-to-strip, silicon-iron, and nickel-iron alloys.

Nickel-Iron Powder Cores. *Nickel-iron powder cores* are made of insulated alloy powder, which is compressed to shape and heat-treated. The alloy composition most widely used is 2-81 Permalloy powder composed of 2% molybdenum, 81% nickel, and balance iron. Another less widely used powder, Sendust, is made of 7% to 13% silicon, 4% to 7% aluminum, and balance iron. Prior to pressing, the powder particles are thinly coated with an inorganic, high-temperature insulation which can withstand the high compacting pressures and the high-temperature (650°C) hydrogen atmosphere anneal. The insulation of the particles lowers eddy-current loss and provides a distributed air gap which can be controlled to provide cores in a range of permeabilities. The 2-81 Permalloy cores are commercially available in permeability ranges of 14 to 300, and Sendust cores have permeabilities ranging from 10 to 140.

These types of nickel-iron powder cores find use in applications where inductance must remain relatively constant when the magnetic component experiences changes in dc current or temperature. Additional stability over temperature can also be achieved by the addition of low-Curie-temperature powder materials to neutralize the naturally positive permeability-temperature coefficient of the alloy powder. Some applications are in telephone loading coils or filter chokes for power conditioning equipment where output voltage ripple must be minimized. Other uses are for pulse transformers and switch-mode power supplies where low power losses are desired. Operating frequencies can range from 1.0 kHz for 300 permeability materials to 500 kHz for the 14 permeability materials.

Powdered-Iron Cores. *Powdered-iron cores* are manufactured from various types of iron powders whose particle sizes range from 2 to 100 μm. The particles are electrically insulated from one another using special insulating materials. The insulated powder is blended with phenolic or epoxy binders and a mold-release agent. The powder is then dry-pressed in a variety of shapes including toroids, E cores, threaded tuning cores, cups, sleeves, slugs, bobbins, and other special shapes. A low-temperature bake of the pressed product produces a solid component in which the insulated particles provide a built-in air gap, reducing eddy-current losses, increasing electrical Q, and thus allowing higher operating frequencies. The use of different iron powder blends and insulation systems provides a range of permeability, from 4 to 90, for use over the frequency spectrum of 50 Hz to 250 MHz. Applications include high-frequency transformers, tuning coils, variable inductors, rf chokes, and noise suppressors for power supply and power control circuits.

Ferrite Cores. *Ferrite cores* are molded from a mixture of metallic oxide powders such that certain iron atoms in the cubic crystal of magnetite (ferrous ferrite) are replaced by other metal atoms, such as Mn and Zn, to form manganese zinc ferrite, or by Ni and Zn to form nickel zinc ferrite. Manganese zinc ferrite is the material most commercially available and is used in devices operating below 1.5 MHz. Nickel zinc ferrites are used mainly for filter applications above that frequency. They resemble ceramic materials in production processes and physical properties. The electrical resistivities correspond to those of semiconductors, being at least 1 million times those of metals. Magnetic permeability μ_0 may be as high as 10,000. The Curie point is quite low, however, in the range 100 to 300°C. Saturation flux density is generally below 5000 G (Fig. 4-12). Ferrite materials are available in several compositions which, through processing, can improve one or two magnetic parameters (magnetic induction, permeability, low hysteresis loss, Curie temperature) at the expense of the other parameters. The materials are fabricated into shapes such as toroids; E, U, and I cores; beads; and self-shielding pot cores. Ferrite cores find use in filter applications up to 1.0 MHz, high-frequency power transformers operating at 10 to 100 kHz, pulse transformer delay lines, adjustable-air-gap inductors, recording heads, and filters used in high-frequency electronic circuits.

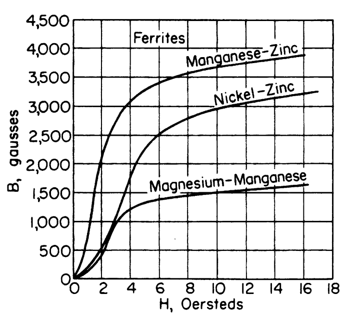


FIGURE 4-12 Typical normal-induction curves for *soft* ferrites.

Permanent-Magnet Materials. *Permanent-magnet materials* that are commercially available may be grouped into five classes as follows:

1. Quench-hardened alloys
2. Precipitation-hardened cast alloys
3. Ceramic materials
4. Powder compacts and elongated single-domain materials
5. Ductile alloys

4.2.10 Quench-Hardened Alloys

Early permanent magnets were made of low-carbon steel (1% C) that was hardened by heat treatment. Later developments saw improvements in the magnetic properties through the use of alloying elements of tungsten, chromium, and cobalt. The chrome steels are less expensive than the cobalt steels, and both find use in hysteresis clutch and motor applications.

Ceramic Magnet Material. *Ceramic magnet material* usage is increasing yearly because of improved magnetic properties and the high cost of cobalt used in metallic alloy magnets. The basic raw material used in these magnets is iron oxide in combination with either strontium carbonate or barium carbonate. The iron oxide and carbonate mixture is calcined, and then the aggregate is ball-milled to a particle size of about 1.0 μm. The material is compacted in dies using the dry powder or a water-based slurry of the powder. High pressures are needed to press the parts to shape. In some ceramic grades, a magnetic field is applied during pressing to orient the material in order to obtain a preferred magnetic orientation. Parts are sintered at high temperatures and ground to finished size using diamond grinding wheels with suitable coolants. Ceramic magnets are hard and brittle, exhibit high electrical resistivities, and have lower densities than cast magnet alloys.

Made in the form of rings, blocks, and arcs, ceramic magnets find use in applications for loudspeakers, dc motors, microwave oven magnetron tubes, traveling wave tubes, holding magnets, chip collectors, and magnetic separator units. Ceramic magnet arcs find wide use in the auto industry in engine coolant pumps, heating-cooling fan motors, and window lift motors. As with other magnets, they are normally supplied nonmagnetized and are magnetized in the end-use structure using magnetizing fields of the order of 10,000 Oe to saturate the magnet. The brittleness of the material necessitates proper design of the magnet support structure so as not to impart mechanical stress to the magnet.

Rare Earth Cobalt Magnets. *Rare earth cobalt magnets* have the highest energy product and coercivity of any commercially available magnetic material. Magnets are produced by powder metallurgy techniques from alloys of cobalt (65% to 77%), rare earth metals (23% to 35%), and sometimes copper and iron. The rare earth metal used is usually samarium, but other metals used are praseodymium, cerium, yttrium, neodymium, lathanum, and a rare earth metal mixture called *misch metal.* The rare earth alloy is ground to a fine particle size (1 to 10 μm), and the powder is then die-compacted in a strong magnetic field. The part is then sintered and abrasive-ground to finish tolerances.

Although this material uses comparatively expensive raw materials, the high value of coercive force (5500 to 9500 Oe) leads to small magnet size and good temperature stability. These magnets find use in miniature electronic devices such as motors, printers, electron beam focusing assemblies, magnetic bearings, and traveling wave tubes. Plastic-bonded rare earth magnets are also being made, but the magnetic value of the energy product is only a fraction of the sintered product.

Ductile Alloys. *Ductile alloys* include the materials Cunife, Vicalloy, Remalloy, chromium-cobalt-iron (Cr-Co-Fe), and in a limited sense, manganese-aluminum-carbon (Mn-Al-C). They are sufficiently ductile and malleable to be drawn, forged, or rolled into wire or strip forms. A final heat treatment after forming develops the magnetic properties. Cunife has a directional magnetism developed as a result of cold working and finds wide use in meters and automotive speedometers. Vicalloy has been used as a high-quality and high-performance magnetic recording tape and in hysteresis clutch applications. Remalloy has been used extensively in telephone receivers but is now being replaced by a newer, less costly magnetic material.

New permanent-magnet materials that are now being produced are the Cr-Co-Fe alloy and the Mn-Al-C alloy. The Cr-Co-Fe alloy family contains 20% to 35% chromium and from 5% to 25% cobalt. This alloy is unique among permanent-magnet alloys due to its good hot and cold ductility, machinability, and excellent magnetic properties. The heat treatment of the alloy involves a rapid cooling from approximately 1200°C to a spinoidal decomposition phase occurring at about 600°C. The magnetic phase developed in the spinoidal decomposition process may be oriented by a heat treatment in a magnetic field, or the material may be magnetically oriented by "deformation aging" as would be accomplished in a wire-drawing operation. The magnetic properties that can be developed are comparable with those of Alnico 5 and are superior to those of the other ductile alloys,

TABLE 4-6 Comparison of Magnetic and Physical Properties of Selected Commercial Materials

	Alnico 5	Alnico 9	Ferrite	Co5R
B_r, G	12800	10500	4100	9500
H_c, Oe	640	1500	2900	6500
B_dH_d, Mg · Oe	5.5	9.0	4.0	22.0
Curie point, °C	850	815	470	740
Temperature coefficient, %/°C	0.02	0.02	0.19	0.03
Density, g/cm^3	7.3	7.3	4.9	8.6
Energy/unit weight	0.8	1.2	0.8	2.6

Cunife, Vicalloy, and Remalloy. Western Electric has introduced a Cr-Co-Fe alloy which replaces Remalloy in the production of telephone receiver magnets and at a lower cost due to reduced cobalt.

The Mn-Al-C Alloy. *The Mn-Al-C alloy* achieves permanent-magnetic properties (B_r, 5500 G; H_c, 2300 Oe; Mg · Oe energy product, 5 Mg · Oe) when mechanical deformation of the alloy takes place at a temperature of about 720°C. Mechanical deformation may be performed by warm extrusion. Magnet size is limited by the amount of deformation needed to develop and orient the magnetic phase in the alloy. The alloying elements are inexpensive, but the tooling and equipment needed in the deformation process is expensive and may be a factor in the economical production of this magnet alloy. Magnets of this alloy would find use in loudspeakers, motor applications, and microwave oven magnetron tubes. The low density, 5.1 g/cm^3, is desirable for motors where reduced inertia and weight savings are important. The low Curie temperature, 320°C, limits the use of this alloy to applications where the ambient temperature is less than 125°C.

Permanent-Magnet Design. *Permanent-magnet design* involves the calculation of magnet area and magnet length to produce a specific magnetic flux density across a known gap, usually with the magnet having the smallest possible volume. Designs are developed from magnet material hysteresis loop data of the second quadrant, commonly called *demagnetization curves*.

Other considerations are the operating temperature of the magnetic assembly, magnet weight, and cost. Also, care should be exercised in the calculation of any steel return path cross section to ensure that it is adequate to carry the flux output of the magnet. Table 4-6 illustrates the range of magnetic characteristics that may be considered in the design. Detailed magnetic and material specifications may be obtained from the magnet manufacturer.

4.3 INSULATING MATERIALS

4.3.1 General Properties

Electrical Insulation and Dielectric Defined. Electrical insulation is a medium or a material which, when placed between conductors at different potentials, permits only a small or negligible current in phase with the applied voltage to flow through it. The term *dielectric* is almost synonymous with electrical insulation, which can be considered the applied dielectric. A perfect dielectric passes no conduction current but only capacitive charging current between conductors. Only a vacuum at low stresses between uncontaminated metal surfaces satisfies this condition.

The range of resistivities of substances which can be considered insulators is from greater than $10^{20}\ \Omega \cdot$ cm downward to the vicinity of $10^6\ \Omega \cdot$ cm, depending on the application and voltage stress. There is no sharp boundary defined between low-resistance insulators and semiconductors. If the voltage stress is low and there is little concern about the level of current flow (other than that which would heat and destroy the insulation), relatively low-resistance insulation can be tolerated.

Circuit Analogy of a Dielectric or Insulation. Any dielectric or electrical insulation can be considered as equivalent to a combination of capacitors and resistors which will duplicate the current-voltage behavior at a particular frequency or time of voltage application. In the case of some dielectrics, simple linear capacitors and resistors do not adequately represent the behavior. Rather, resistors and capacitors with particular nonlinear voltage-current or voltage-charge relations must be postulated to duplicate the dielectric current-voltage characteristic.

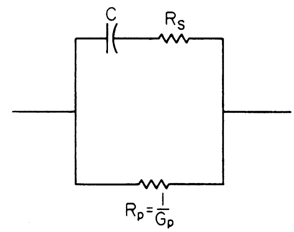


FIGURE 4-13 Equivalent circuit of a dielectric.

The simplest circuit representation of a dielectric is a parallel capacitor and resistor, as shown in Fig. 4-13 for $R_S = 0$. The perfect dielectric would be simply a capacitor. Another representation of a dielectric is a series-connected capacitor and resistor as in Fig. 4-13 for $R_p = \infty$, while still another involves both R_S and R_p.

The ac dielectric behavior is indicated by the phase diagram (Fig. 4-14). The perfect dielectric capacitor has a current which leads the voltage by 90°, but the imperfect dielectric has a current which leads the voltage by less than 90°. The dielectric phase angle is θ, and the difference, $90° - \theta = \delta$, is the loss angle. Most measurements of dielectrics give directly the tangent of the loss angle tan δ (known as the *dissipation factor*) and the capacitance C. In Fig. 4-13, if $R_p = \infty$, the series $R_s - C$ has a tan $\delta = 2\pi f C_s R_s$, and if $R_s = 0$, the parallel $R_p - C$ has a tan $\delta = 1/2\pi f C_p R_p$.

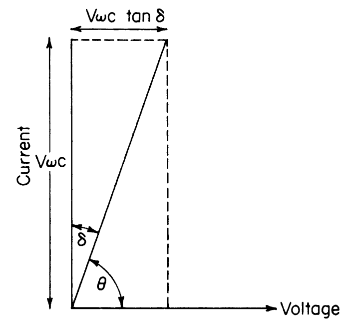


FIGURE 4-14 Current-voltage phase relation in a dielectric.

The ac power or heat loss in the dielectric is $V^2 2\pi f C$ tan δ watts, or VI sin δ watts, where sin δ is known as the power factor, V is the applied voltage, I is the total current through the dielectric, and f is the frequency. From this it can be seen that the equivalent parallel conductance of the dielectric σ (the inverse of the equivalent parallel resistance ρ) is $2\pi f C$ tan δ. The ac conductivity is

$$\sigma = (5/9) f \epsilon' \tan \delta \times 10^{-12}\ \Omega^{-1}\ \text{cm}^{-1} = 1/\rho \tag{4-54}$$

where ϵ' is the permittivity (or relative dielectric constant) and f is the frequency. (The IEEE now recommends the symbol ϵ' for the dielectric constant relative to a vacuum. The literature on dielectrics and insulation also has used κ [kappa] for this dimensionless quantity or ϵ'_r. In some places, ϵ' has been used to indicate the absolute dielectric constant, which is the product of the relative dielectric constant and the dielectric constant of a vacuum ϵ_0, which is equal to 8.85×10^{-12} F/m.) κ_0 also has been used to represent the dielectric constant of a vacuum. While the ac conductivity theoretically increases in proportion to the frequency, in practice, it will depart from this proportionality insofar as ϵ' and tan δ change with frequency.

Capacitance and Permittivity or Dielectric Constant. The capacitance between plane electrodes in a vacuum (with fringing neglected) is

$$C = \epsilon' \epsilon_0 A/t = 0.0884 \times 10^{-12} A/t \qquad \text{farads} \tag{4-55}$$

where ϵ_0 is the dielectric constant of a vacuum, A is the area in square centimeters, and t is the spacing of the plates in centimeters. ϵ_0 is 0.225×10^{-12} F/in when A and t are expressed in inch units.

When a dielectric material fills the volume between the electrodes, the capacitance is higher by virtue of the charges within the molecules and atoms of the material, which attract more charge to the capacitor planes for the same applied voltage. The capacitance with the dielectric between the electrodes is

$$C = \epsilon' \epsilon_0 A/t \tag{4-56}$$

where ϵ' is the relative dielectric constant of the material. The capacitance relations for several other commonly occurring situations are

Coaxial conductors: $$C = \frac{2\pi\epsilon'\epsilon_0 L}{\ln(r_2/r_1)} \quad \text{farads} \tag{4-57}$$

Concentric spheres: $$C = \frac{4\pi\epsilon'\epsilon_0 r_1 r_2}{r_2 - r_1} \quad \text{farads} \tag{4-58}$$

Parallel cylindrical conductors: $$C = \frac{2\pi\epsilon'\epsilon_0 L}{\cosh^{-1}(D/2r)} \quad \text{farads} \tag{4-59}$$

In these equations, L is the length of the conductors, r_2 and r_1 are the outer and inner radii, and D is the separation between centers of the parallel conductors with radii r. For dimensions in centimeters, ϵ_0 is 0.0884 F/cm.

The value of ϵ' depends on the number of atoms or molecules per unit volume and the ability of each to be polarized (i.e., to have a net displacement of their charge in the direction of the applied voltage stress). Values of ϵ' range from unity for vacuum to slightly greater than unity for gases at atmospheric pressure, 2 to 8 for common insulating solids and liquids, 35 for ethyl alcohol and 91 for pure water, and 1000 to 10,000 for titanate ceramics (see Table 4-7 for typical values).

The relative dielectric constant of materials is not constant with temperature, frequency, and many other conditions and is more appropriately called the *dielectric permittivity*. Refer to the volume by Smyth (1955) for a discussion of the relation of ϵ' to molecular structure and to von Hippel (1954) and other tables of dielectric materials from the MIT Laboratory for Insulation Research. The permittivity of many liquids has been tabulated in *NBS Circ.* 514. The *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Co.) also lists values for a number of plastics and other materials.

The permittivity of many plastics, ceramics, and glasses varies with the composition, which is frequently variable in nominally identical materials. In the case of some plastics, it varies with degree of cure and in the case of ceramics with the firing conditions. Plasticizers often have a profound effect in raising the permittivity of plastic compositions.

There is a force of attraction between the plates of a capacitor having an applied voltage. The stored energy is $\frac{1}{2} CV^2$ J. The force equals the derivative of this energy with respect to the plate separation: $(\frac{1}{2})\, \epsilon'\epsilon_0 E^2 \times 10^2$ N/cm² or $(\frac{1}{2})\, \epsilon'\epsilon_0 E^2 \times 10$ bar, where E is the electric field in volts per centimeter. The force increases proportionally to the capacitance or permittivity. This leads to a force of attraction of dielectrics into an electric field, that is, a net force which tends to move them toward a region of high field. If two dielectrics are present, the one with higher permittivity will displace the one with lower permittivity in the higher-field region. For example, air bubbles in a liquid are repelled from high-field regions. Correspondingly, elongated dielectric bodies are rotated into the direction of the electric field. In general, if the voltage on a dielectric system is maintained constant, the dielectrics move (if they are able) to create a higher capacitance.

Resistance and Resistivity of Dielectrics and Insulation. The measured resistance R of insulation depends on the geometry of the specimen or system measured, which for a parallel-plate arrangement is

$$R = \rho t/A \quad \text{ohms} \tag{4-60}$$

where t is the insulation thickness in centimeters, A is the area in square centimeters, and ρ is the dielectric resistivity in ohm-centimeters. If t and A vary from place to place, the effective "insulation resistance" will be determined by the effective integral of the t/A ratio over all the area under stress, on the assumption that the material resistivity ρ does not change. If the material is not homogeneous and materials of different resistivities appear in parallel, the system can be treated as parallel

TABLE 4-7 Dielectric Permittivity (Relative Dielectric Constant), ϵ'

	k		k
Inorganic crystalline		Polymer resins	
NaCl, dry crystal	5.5	Nonpolar resins	
$CaCO_2$ (av)	9.15	Polyethylene	2.3
Al_2O_2	10.0	Polystyrene	2.5–2.6
MgO	8.2	Polypropylene	2.2
BN	4.15	Polytetrafluoroethylene	2.0
TiO_2 (av)	100	Polar resins	
$BaTiO_2$ crystal	4,100	Polyvinyl chloride (rigid)	3.2–3.6
Muscovite mica	7.0–7.3	Polyvinyl acetate	3.2
Fluorophlogopite (synthetic mica)	6.3	Polyvinyl fluoride	8.5
		Nylon	4.0–4.6
Ceramics		Polyethylene terephthalate	3.25
Alumina	8.1–9.5	Cellulose cotton fiber (dry)	5.4
Steatite	5.5–7.0	Cellulose Kraft fiber (dry)	5.9
Forsterite	6.2–6.3	Cellulose cellophane (dry)	6.6
Aluminum silicate	4.8	Cellulose triacetate	4.7
Typical high-tension porcelain	6.0–8.0	Tricyanoethyl cellulose	15.2
Titanates	50–10,000	Epoxy resins unfilled	3.0–4.5
Beryl	4.5	Methylmethacrylate	3.6
Zirconia	8.0–10.5	Polyvinyl acetate	3.7–3.8
Magnesia	8.2	Polycarbonate	2.9–3.0
Glass-bonded mica	6.4–9.2	Phenolics (cellulose-filled)	4–15
		Phenolica (glass-filled)	5–7
Glasses		Phenolics (mica-filled)	4.7–7.5
Fused silica	3.8	Silicones (glass-filled)	3.1–4.5
Corning 7740 (common laboratory Pyrex)	5.1		

resistors: $R = R_aR_b/(R_a + R_b)$. In this case, the lower-resistivity material usually controls the overall behavior. But if materials of different resistivities appear in series in the electric field, the higher-resistivity material generally will control the current, and a majority of the voltage will appear across it, as in the case of series resistors.

The resistance of dielectrics and insulation is usually time-dependent and (for the same reason) frequency-dependent. The dc behavior of dielectrics under stress is an extension of the low-frequency behavior. The ac and dc resistance and permittivity can, in principle, be related for comparable times and frequencies.

Current flow in dielectrics can be divided into parts: (*a*) the true dc current, which is constant with time and would flow indefinitely, is associated with a transport of charge from one electrode into the dielectric, through the dielectric, and out into the other electrode, and (*b*) the polarization or absorption current, which involves, not charge flow through the interface between the dielectric and the electrode, but rather the displacement of charge within the dielectric. This is illustrated in Fig. 4-15, where it is shown that the displaced or absorbed charge is responsible for a reverse current when the voltage is removed.

Polarization current results from any of the various forms of limited charge displacement which can occur in the dielectric. The displacement occurring first (within less than nanoseconds) is the electronic and intramolecular charged atom displacement responsible for the very high frequency permittivity.

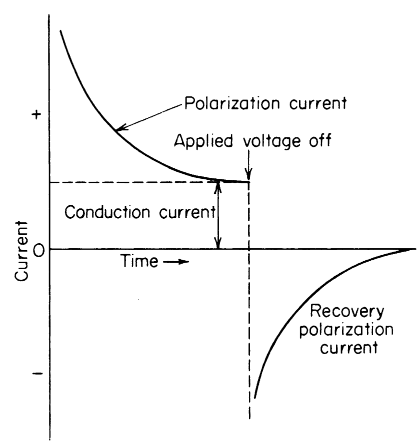


FIGURE 4-15 Typical dc dielectric current behavior.

The next slower displacement is the rotation of dipolar molecules and groups which are relatively free to move. The displacement most commonly observed in dc measurements, that is, currents changing in times of the order of seconds and minutes, is due to the very slow rotation of dipolar molecules and ions moving up to internal barriers in the material or at the conductor surfaces. When those slower displacement polarizations occur, the dielectric constant declines with increasing frequency and approaches the square of the optical refractive index η^2 at optical frequencies.

In composite dielectrics (material with relatively lower resistance intermingled with a material of relatively higher resistance), a large interfacial or Maxwell-Wagner type of polarization can occur. A circuit model of such a situation can be represented by placing two of the circuits of Fig. 4-13 in series and making the parallel resistance of one much lower than the other. To get the effect, it is necessary that the time constant R_pC be different for each material.

A simple model of the polarization current predicts an exponential decline of the current with time: $I_p = Ae^{-\alpha t}$, similar to the charging of a capacitor through a resistor. Composite materials are likely to have many different time constants, $\alpha = 1/RC$, superimposed. It is found empirically that the polarization or absorption current decreases inversely as a simple negative exponent of the time

$$I = At^{-n} \tag{4-61}$$

The ratio of the current at 1 min to that at 10 min has been called the *polarization index* and is used to indicate the quality of composite machine insulation. A low polarization index associated with a low resistance sometimes indicates parallel current leakage paths through or over the surface of insulation (e.g., in adsorbed water films).

The level of the conduction current which flows essentially continuously through insulation is an indication of the level of the ionic concentration and mobility in the material. Frequently, as with salt in water, the ions are provided by dissolved, absorbed, or included impurity electrolytes in the material rather than by the material itself. Purifying the material will therefore often raise the resistivity. If it is liquid, purification can be done with adsorbent clays or ion-exchange resins.

The conductivity of ions in an insulation is given by the equation

$$\sigma = \mu ec \qquad \Omega^{-1}\cdot\text{cm}^{-1} \tag{4-62}$$

where μ is the ion mobility, e is the ionic charge in coulombs, and c is the ionic concentration per cubic centimeter. The mobility, expressed in centimeters per second-volt per centimeter, decreases inversely with the effective internal viscosity and is very low for hard resins, but it increases with temperature and with softness of the resin or fluidity of liquids. The ionic conductivity also varies widely with material purity. Among the polymers and resins, nonpolar resins such as polyethylene are likely to have high resistivities, on the order of 10^{16} or greater, since they do not readily dissolve or dissociate ionic impurities. Harder or crystalline polar resins have higher resistivity than do similar softer resins of similar dielectric constant and purity. Resins and liquids of higher dielectric constant usually have higher conductivities because they dissolve ionic impurities better, and the impurities dissociate to ions much more readily in a higher dielectric constant medium. Ceramics and glasses have lower resistivity if they contain alkali ions (sodium and potassium), since these ions are highly mobile.

Water is particularly effective in decreasing the resistivity by increasing the ionic concentration and mobility of materials, on the surface as well as internally. Water associates with impurity ions or ionizable constituents within or on the surface or interfaces. It helps to dissociate the ions by virtue of its high dielectric constant and provides a local environment of greater mobility, particularly as surface water films.

The ionic conductivity σ, exclusive of polarization effects, can be expected to increase exponentially with temperature according to the relation

$$\sigma = \sigma_0 e^{-B/T} \tag{4-63}$$

where T is the Kelvin temperature and σ_0 and B are constants. This relation, log σ versus $1/T$, is shown in Fig. 4-16. It is often observed that at lower temperatures, where the resistivity is higher, the resistivity tends to be lower than the extrapolated higher temperature line would predict. There are at least two possible reasons for this: the effect of adsorbed moisture and the contribution of a very slowly decaying polarization current.

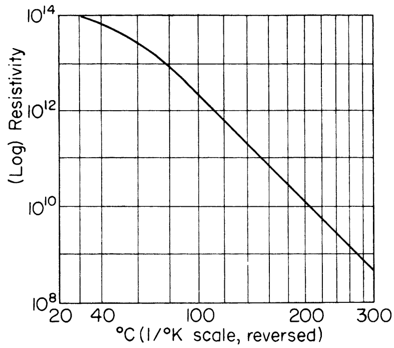


FIGURE 4-16 Typical dielectric resistivity-temperature dependence. (*Corning*.)

Variation of Dielectric Properties with Frequency. The permittivity of dielectrics invariably tends downward with increasing frequency, owing to the inability of the polarizing charges to move with sufficient speed to follow the increasing rate of alternations of the electric field. This is indicated in Fig. 4-17. The sharper decline in permittivity is known as a *dispersion region.* At the lower frequencies, the ionic-interface polarization declines first; next, the molecular dipolar polarizations decline. With some polar polymers, two or more dipolar dispersion regions may occur owing to different parts of the molecular rotation.

Figure 4-17 is typical of polymers and liquids but not of glasses and ceramics. Glasses, ceramics, and inorganic crystals usually have much flatter permittivity-frequency curves, similar to that shown for the nonpolar polymer, but at a higher level, owing to their atom-ion displacement polarization, which can follow the electric field usually up to infrared frequencies.

The dissipation factor–frequency curve indicates the effect of ionic migration conduction at low frequency. It shows a maximum at a frequency corresponding to the permittivity dispersion region. This maximum is usually associated with a molecular dipolar rotation and occurs when the rotational mobility is such that the molecular rotation can just keep up with frequency of the applied field. Here it has its maximum movement in phase with the voltage, thus contributing to conduction current. At lower frequencies, the molecule dipole can rotate faster than the field and contributes more to permittivity. At higher frequencies it cannot move fast enough. Such a dispersion region can also occur because of ionic migration and interface polarization if the interfaces are closely spaced and if the frequency and mobility have the required values.

The frequency region where the dipolar dispersion occurs depends on the rotational mobility. In mobile, low-viscosity liquids, it is in the 100- to 10,000-MHz range. In viscous liquids, it occurs in the region of 1 to 100 MHz. In soft polymers it may occur in the audio-frequency range, and with hard polymers it is likely to be at very low frequency (indistinguishable from dc properties). Since the viscosity is affected by the temperature, increased temperature shifts the dispersion to higher frequencies.

Variation of Dielectric Properties with Temperature. The trend in ac permittivity and conductivity, as measured by the dissipation factor, is controlled by the increasing ionic migrational and dipolar molecular rotational mobility with increasing temperature. This curve, which is indicated

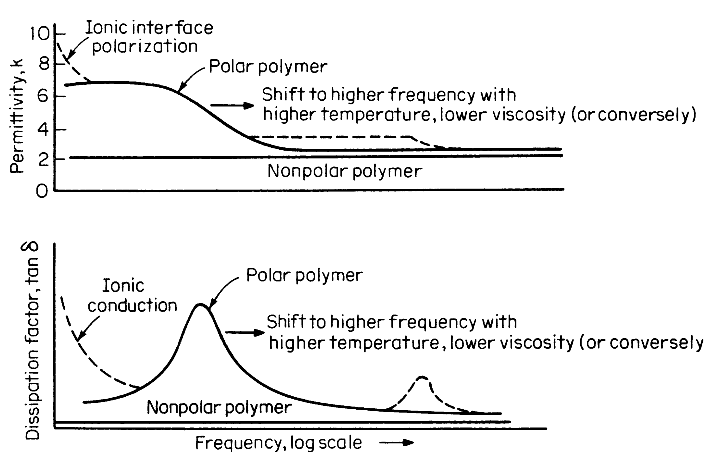


FIGURE 4-17 Typical variation in dielectric properties with frequency.

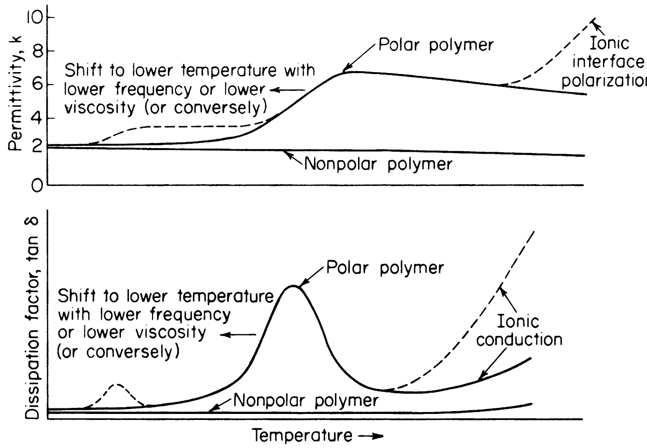


FIGURE 4-18 Typical variation in dielectric properties with temperature.

in Fig. 4-18, is in most respects a mirror image of the frequency trend shown in Fig. 4-17, since the two effects are interrelated.

The permittivity-dispersion and dissipation factor maximum region occurs below room temperature for viscous liquids and still lower for mobile liquids. In fact, mobile liquids may crystallize before they would show dispersion, except at high frequencies. With polymers, the dissipation factor maxima is likely to occur, at power frequencies, at a temperature close to a softening-point or internal second-order transition-point temperature. Dielectric dispersion and mechanical modulus dispersion usually can be correlated at the same temperature for comparable frequencies.

Composite Dielectrics. The dielectric properties of composite dielectrics are generally a weighted average of the individual component properties, unless there is interaction, such as dissolving (as opposed to intermixing) of one material in another, or chemical reaction of one with another. Interfaces created by the mixing present a special factor, which often can lead to a higher dissipation factor and lower resistivity as a result of moisture and/or impurity concentration at the interface.

The ac properties of sheets of two dielectrics of dielectric constant k_1 and k_2 and of thickness t_1 and t_2 placed in series are related to the properties of the individual materials by the series of capacitance and impedance relation

$$C = \frac{k_0 k_1 k_2 A}{k_1 t_2 + k_2 t_1} \tag{4-64}$$

$$\tan\delta = \frac{(t_1/t_2)\epsilon'_2 \tan\delta_1 + \epsilon'_1 \tan\delta_2}{\epsilon'_1 + \epsilon'_2 (t_1/t_2)} \tag{4-65}$$

Similarly, the properties of two dielectrics in parallel are

$$C = \epsilon_0 \left(\frac{\epsilon'_1 A_1}{t_1} + \frac{\epsilon'_2 A_2}{t_2} \right) \tag{4-66}$$

$$\tan\delta = \frac{t_2 \epsilon'_1 A_1 \tan\delta_1 + t_1 \epsilon'_2 A_2 \tan\delta_2}{t_2 \epsilon'_1 A_1 + t_1 \epsilon'_2 A_2} \tag{4-67}$$

With steady dc voltages, the resistivities control the current. With equal-area layer dielectrics in series,

$$R = R_1 + R_2 = \frac{1}{A}(\rho_1 t_1 + \rho_2 t_2) \tag{4-68}$$

When the dielectrics are in parallel and of equal thickness t,

$$R = \frac{R_1 R_2}{R_1 + R_2} = \frac{\rho_1 \rho_2 t}{\rho_1 A_2 + \rho_2 A_1} \tag{4-69}$$

Potential Distribution in Dielectrics. The maximum potential gradient in dielectrics is of critical significance insofar as the breakdown is concerned, since breakdown or corona is usually initiated at the region of highest gradient. In a uniform-field arrangement of conductors or electrodes, the maximum gradient is simply the applied voltage divided by the minimum spacing. In divergent fields, the gradient must be obtained by calculation (which is possible for some simple arrangements) or by field mapping.

A common situation is the coaxial geometry with inner and outer radii R_1 and R_2. The gradient at radius r (centimeters) with voltage V applied is given by the equation

$$E = \frac{V}{r \ln (R_2/R_1)} \qquad \text{V/cm} \tag{4-70}$$

The gradient is a maximum at $r = R_1$.

When different dielectrics appear in series, the greater stress with ac fields is on the material having the lower dielectric constant. This material will frequently break down first unless its dielectric strength is much higher

$$\frac{E_1}{E_2} = \frac{\epsilon'_2}{\epsilon'_1} \quad \text{and} \quad E_1 = \frac{V}{t_1 + t_2 \epsilon'_1 / \epsilon'_2} \tag{4-71}$$

The effect of the insulation thickness and dielectric constant (as well as the sharpness of the conductor edge) to create sufficient electric stress for local air breakdown (partial discharges) is shown in Fig. 4-19. With dc fields, the stress distributes according to the resistivities of the materials, the higher stress being on the higher-resistivity material.

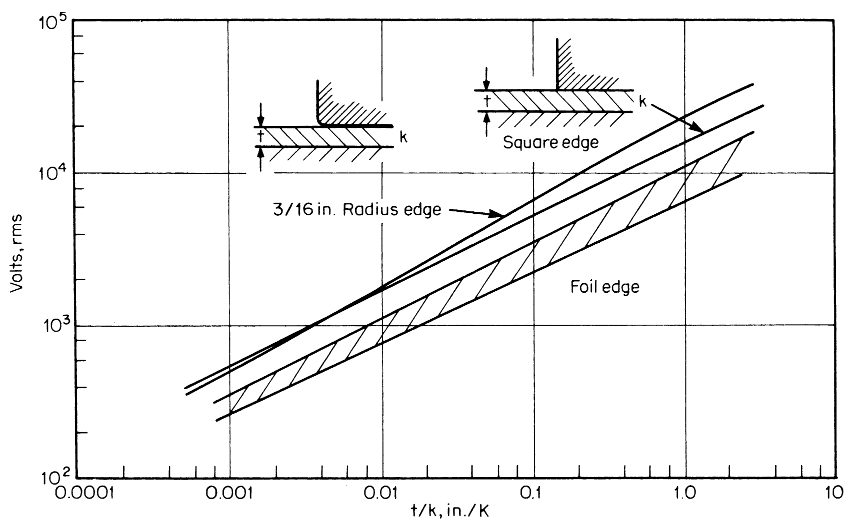


FIGURE 4-19 Corona threshold voltage at conductor edges in air as a function of insulation thickness.

Dielectric Strength. This is defined by the ASA as the maximum potential gradient that the material can withstand without rupture. Practically, the strength is often reported as the breakdown voltage divided by the thickness between electrodes, regardless of electrode stress concentration.

Breakdown appears to require not only sufficient electric stress but also a certain minimum amount of energy. It is a property which varies with many factors such as thickness of the specimen, size and shape of electrodes used in applying stress, form or distribution of the field of electric stress in the material, frequency of the applied voltage, rate and duration of voltage application, fatigue with repeated voltage applications, temperature, moisture content, and possible chemical changes under stress.

The practical dielectric strength is decreased by defects in the material, such as cracks, and included conducting particles and gas cavities. As will be shown in more detail in later sections on gases and liquids, the dielectric strength is quite adversely affected by conducting particles.

To state the dielectric strength correctly, the size and shape of specimen, method of test, temperature, manner of applying voltage, and other attendant conditions should be particularized as definitely as possible.

ASTM standard methods of dielectric strength testing should be used for making comparison tests of materials, but the levels of dielectric strength measured in such tests should not be expected to apply in service for long times. It is best to test an insulation in the same configuration in which it would be used. Also, the possible decline in dielectric strength during long-time exposure to the service environment, thermal aging, and partial discharges (corona), if they exist at the applied service voltage, should be considered. ASTM has thermal life test methods for assessing the long-time endurance of some forms of insulation such as sheet insulation, wire enamel, and others. There are IEEE thermal life tests for some systems such as random wound motor coils.

The dielectric strength varies as the time and manner of voltage application. With unidirectional pulses of voltage, having rise times of less than a few microseconds, there is a time lag of breakdown, which results in an apparent higher strength for very short pulses. In testing sheet insulation in mineral oil, usually a higher strength for pulses of slow rise time and somewhat higher strength for dc voltages is observed.

The trend in breakdown voltage with time is typical of many solid insulation systems.

With ac voltages, the apparent strength declines steadily with time as a result of partial discharges (in the ambient medium at the conductor or electrode edge). These penetrate the solid insulation. The discharges result from breakdown of the gas or liquid prior to the breakdown of the solid. Mica in particular, as well as other inorganic materials, is more resistant to such discharges. Organic resins should be used with caution where the ac voltage gradient is high and partial discharges (corona) may be present. Since the presence of partial discharges on insulation is so important to the long-time voltage endurance, their detection and measurement have become very important quality control and design tools. If discharges continuously strike the insulation within internal cavities or on the surface, the time to failure usually varies inversely as the applied frequency, since the number of discharges per unit time increases almost in direct proportion to the frequency. But in some cases, ambient conditions prevent continuous discharges.

When organic resin insulation is fabricated to avoid partial discharges using conductors or electrodes intimately bonded to the insulation, as in extruded polyethylene cables with a plastic semiconducting interface between the resin and the coaxial inner and outer metal conductors, respectively, the voltage endurance is greatly extended. Imperfections, however, in this "semicon"-resin interface, or at conducting particle inclusions in the resin, can lead to local discharges and the development of "electrical tree" growth. Vacuum impregnating and casting electrodes or conductors into resin also tend to avoid cavities and surface discharges and greatly improve the voltage endurance at high stresses.

The dc strength of solid insulation is usually higher and declines much less with time than the ac strength, since corona discharges are infrequent.

The dielectric strength is much higher where surface discharges are avoided and when the electric field is uniform. This can be achieved with solid materials by recessing spherical cavities into the material and using conducting paint electrodes.

The "intrinsic" electric strength of solid materials measured in uniform fields, avoiding surface discharges, ranges from levels on the order of 0.5 to 1 MV/cm for alkali halide crystals, which are about the lowest, upward to somewhat more than 10 MV/cm. Polymers and some inorganic materials, such as mica and aluminum oxide, have strengths of 2 to 20 MV/cm for thin films. The strength decreases with increasing thickness and with temperature above a critical temperature (which is usually from 1 to 100°C), below which the strength has a level value or a moderate increase with increasing temperature. Below the critical temperature, the breakdown is believed to be strictly electronic in nature and is constant or increases slightly with temperature. Above this temperature, it declines owing to dielectric thermal heating.

The breakdown voltage of thin insulation materials containing defects, which give the minimum breakdown voltage, declines as the area under stress increases. The effect of area on the strength can be estimated from the standard deviation S of tests on smaller areas by applying minimum value statistics: $V_1 - V_2 = 1.497\, S \log(A_1/A_2)$, where V_1 and V_2 are the breakdown voltages of areas A_1 and A_2.

If the ac or dc conductivity of a dielectric is high or the frequency is high, breakdown can occur as a result of dielectric heating, which raises the temperature of the material sufficiently to cause melting or decomposition, formation of gas, etc. This effect can be detected by measuring the conductivity as a function of applied electric stress. If the conductivity rises with time, with constant voltage, and at constant ambient temperature, this is evidence of an internal dielectric heating. If the heat transfer to the electrodes and ambient surroundings is adequate, the internal temperature eventually may stabilize, but if this heat transfer is inadequate, the temperature will rise until breakdown occurs. The criterion of this sort of breakdown is the heat balance between dielectric heat input and loss to the surroundings.

The dielectric heat input is given by the equation

$$\sigma E^2 = (5/9\, \epsilon' f \tan \delta \times 10^{-12}) E^2 \qquad \text{W/cm}^3 \tag{4-72}$$

where E is the field in volts per centimeter. When this quantity is on the order of 0.1 or greater, dielectric heating can be a problem. It is much more likely to occur with thick insulation and at elevated temperatures.

Water Penetration. Water penetration into electrical insulation also degrades the dielectric strength by several mechanisms. The effect of water to increase the insulation conductivity contributes thereby to a decreased dielectric strength, probably by a thermal breakdown mechanism. Another effect noticed recently, particularly in polyethylene cables, is the development of "water" or "electrochemical trees." Water (and/or a similar high dielectric constant chemical) can diffuse through polyethylene and collect at tiny hygroscopic inclusion sites, where the water or chemical is adsorbed. Then the electric field causes an expansion and growth of the adsorbed water or chemical in the electric field direction. This may completely bridge the insulation or possibly increase the local electric stress at the site so as to produce an electric tree and eventual breakdown.

Ionizing Radiation. Ionizing radiation, as from nuclear sources, may degrade insulation dielectric strength and integrity by causing polymer chain scission, and cracking of some plastics, as well as gas bubbles in liquids. Also, the conductivity levels in solids and liquids are increased.

Arc Tracking of Insulation. High-current arc discharges between conductors across the surface of organic resin insulation may carbonize the material and produce a conducting track. In the presence of surface water films, formed from rain or condensation, etc., small arc discharges form between interrupted parts of the water film, which is fairly conducting, and conducting tracks grow progressively across the surface, eventually bridging between conductors and causing complete breakdown. Materials vary widely in their resistance to tracking, and there are a variety of dry and wet tests for this property. With proper fillers, some organic resins can be made essentially nontracking. Some resins such as polymethyl methacrylate and polymethylene oxide burst into flame under arcing conditions.

Thermal Aging. Organic resinous insulating materials in particular are subject in varying degrees to deterioration due to thermal aging, which is a chemical process involving decomposition or modification of the material to such an extent that it may no longer function adequately as the intended insulation. The aging effects are usually accelerated by increased temperature, and this characteristic is used to make accelerated tests to failure or to an extent of deterioration considered dangerous. Such tests are made at appreciably higher than normal operating temperatures, if the expected life is to be several years or more, since useful accelerated tests reasonably should be completed in less than a year.

Frequently, other environmental factors influence the life in addition to the temperature. These include presence or absence of oxygen, moisture, and electrolysis. Mechanical and electrical stress may reduce the life by setting a required level of performance at which the insulation must perform. If this level is high, less deterioration of the insulation is required to reach this level.

Sometimes a complete apparatus is life-tested, as well as smaller specimens involving only one insulation material or a simple combination of these in a simple model. New tests are being devised continually, but there has been some standardization of tests by the IEEE and ASTM and internationally by the IEC.

It is important to note that frequently materials are assigned temperature ratings based on tests of the material alone. Often that material, combined with others in an apparatus or system, will perform satisfactorily at appreciably higher temperatures. Conversely, because of incompatibility with other materials, it may not perform at as high a temperature as it would alone. For this reason, it is considered desirable to make functional operating tests on complete systems. These can also be accelerated at elevated temperatures and environmental exposure conditions such as humidification, vibration, cold-temperature cycling, etc. introduced intermittently. The basis for temperature rating of apparatus and materials is discussed thoroughly in *IEEE Standard Publ.* 1. Tests for determining ratings are described in *IEEE Publs.* 98, 99, and 101.

Application of Electrical Insulation. In applying an insulating material, it is necessary to consider not only the electrical requirements but also the mechanical and environmental conditions of the application. Mechanical failure often leads to electrical failure, and mechanical failure is frequently the primary cause for failure of an aged insulation.

The initial properties of an insulation are frequently more than adequate for the application, but the effects of aging and environment may degrade the insulation rapidly to the point of failure. Thus, the thermal and environmental stability should be considered of equal importance. The effects of moisture and surface dirt contamination should be particularly considered, if these are likely to occur.

4.3.2 Insulating Gases

General Properties of Gases. A gas is a highly compressible dielectric medium, usually of low conductivity and with a dielectric constant only a little greater than unity, except at high pressures. In high electric fields, the gas may become conducting as a result of impact ionization of the gas molecules by electrons accelerated by the field and by secondary processes which produce partial breakdown (corona) or complete breakdown. Conditions which ionize the gas molecules, such as very high temperatures and ionizing radiation (ultraviolet rays, x-rays, gamma rays, high-velocity electrons, and ions such as alpha particles), will also produce some conduction in a gas.

The gas density d (grams per liter) increases with pressure p (torrs or millimeters of mercury) and gram-molecular weight M and decreases inversely with the absolute temperature T (degrees Celsius + 273) according to the relation

$$d = \frac{M}{22.4}\frac{p}{760}\frac{273}{T} \qquad \text{g/L} \tag{4-73}$$

The preceding relation is exact for ideal gases but is only approximately correct for most common gases.

If the gas is a vapor in equilibrium with a liquid or solid, the pressure will be the vapor pressure of the liquid or solid. The logarithm of the pressure varies as $-\Delta H/RT$, where ΔH is the heat of vaporization in calories per mole and R is the molar gas constant, 1.98 cal/(mol)(°C). This relation also applies to all common atmospheric gases at low temperatures, below the points where they liquify.

Dielectric Properties at Low Electric Fields

Dielectric Constant. The dielectric constant k of gases is a function of the molecular electrical polarizability and the gas density. It is independent of magnetic and electric fields except when a significant number of ions is present.

Conduction. The conductivity of a pure molecular gas at moderate electric stress and moderate temperature can be assumed, in the absence of any ionizing effect such as ionizing radiation, to be practically zero. Ionizing radiation induces conduction in the gas to a significant extent, depending on the amount absorbed and the volume of gas under stress. The energy of the radiation must exceed, directly or indirectly, the ionization energy of the gas molecules and thus produce an ion pair (usually an electron and positive ion). The threshold ionization energy is on the order of 10 to 25 electronvolts (eV)/molecule for common gases (10.86 eV for methyl alcohol, 12.2 for oxygen, 15.5 for nitrogen, and 24.5 for helium). Only very short wavelength ultraviolet light is effective directly in photoionization, since 10 eV corresponds to a photon of ultraviolet with a wavelength of 1240 Å. Since the photoelectric work function of metal surfaces is much lower (2 to 6 eV; e.g., copper about 4 eV), the longer-wavelength ultraviolet commonly present is effective in ejecting electrons from a negative conductors surface. Such cathode-ejected electrons give the gas apparent conductivity.

High-energy radiation from nuclear disintegration is a common source of ionization in gases. Nuclear sources usually produce gamma rays on the order of 10^6 eV energy. Only a small amount is absorbed in passing through a low-density gas. A flux of 1 R/h produces ion pairs corresponding to a saturation current (segment *ab* of Fig. 4-20) of 0.925×10^{-13}A/cm^3 of air at 1 atm pressure if all the ions formed are collected at the electrodes. The effect is proportional to the flux and the gas density.

At a voltage stress below about 100 V/cm, some of the ions formed will recombine before being collected, and the current will be correspondingly less (segment *oa* of Fig. 4-20). Higher stresses do not increase the current if all the ions formed are collected. A very small current, on the order of 10^{-21}A/cm^3 of air, is attributable to cosmic rays and residual natural radioactivity.

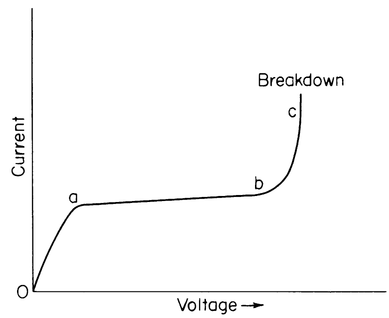


FIGURE 4-20 Current-voltage behavior of a lightly ionized gas.

Electrons (beta rays) produce much more ionization per path length than gamma rays, because they are slowed down by collisions and lose their energy more quickly. Correspondingly, the slower alpha particles (positive helium nuclei) produce a very dense ionization in air over a short range. For example, a 3-million-eV (MeV) alpha particle has a range in air of 1.7 cm and creates a total of 6.8×10^5 ion pairs. A beta particle (an electron) of the same energy creates only 40 ion pairs per centimeter and has a range of 13 m in air.

It should be noted that ionizing radiation of significant levels has only a small effect on gas dielectric strength. For example, the ionization current produced by a corona discharge from a needle point is typically much higher than that produced by a radiation flux of significant level, 10^{11} gamma photons per square centimeter.

At temperatures increasing above 600°C, it has been shown that thermionic electron emissions from negative conductor surfaces produce significant currents compared with levels typical of electrical insulation.

Since the rate of production of ions by the various sources mentioned above is limited, the current in the gas does not follow Ohm's law, unless the rate of collection of the ions at the electrodes is small compared with the rate of production of these ions, as in the initial part of segment *oa* in Fig. 4-20.

Dielectric Breakdown

Uniform Fields. The dielectric breakdown of gases is a result of an exponential multiplication of free electrons induced by the field. It is generally assumed that the initiation of breakdown

TABLE 4-8 Relative Dielectric Strengths of Gases (0.1-in gap)

Air	0.95	CF_4	1.1
N_2	1.0	C_2F_6	1.9
CO_2	0.90	C_3F_8	2.3
H_2	0.57	C_4F_8 cyclic	2.8
A	0.28	CF_2Cl_2	2.4
Ne	0.13	C_2F_5Cl	2.6
He	0.14	$C_2F_4Cl_2$	3.3
SF_6	2.3–2.5		

requires only one electron. However, if only a few electrons are present prior to breakdown, it is not easily possible to measure the trend of current shown in Fig. 4-20. If the breakdown is completed between metal electrodes, the spark develops extremely rapidly into an arc, involving copious emission of electrons from the cathode metal and, if the necessary current flow is permitted, vaporization of metal from the electrodes. Table 4-8 gives the dielectric strength of typical gases.

In uniform electric fields, breakdown occurs at a critical voltage which is a function of the product of the pressure p and spacing d (Paschen's law).

It would be more accurate to consider the gas density-spacing product, since the dielectric strength varies with the temperature only as the latter affects the gas density. It will be noted that the electric field at breakdown decreases as the spacing increases. This is typical of all gases and is due to the fact that a minimum amount of multiplication of electrons must occur before breakdown occurs. A single electron accelerated by the field creates an avalanche which grows exponentially as $e^{\alpha}x$, where x is the distance and α is the Townsend ionization coefficient (electrons formed by collision per centimeter), which increases rapidly with electric field. At small spacings, α and the field must be higher for sufficient multiplication. In divergent electric fields or large spacings, it has been found that when the integral *int*$\alpha_E\,dx$ increases to about 18.4 (10^8 electrons), sufficient space charge develops to produce a streamer type of breakdown. It seems to be apparent that the final step in gas breakdown before arc development is the development of a branched filamentary streamer which proceeds more easily from the positive electrode toward the negative electrode.

Relative Dielectric Strengths of Gases. The relative dielectric strength, with few exceptions, tends upward with increasing molecular weight. There are a number of factors other than molecular or atomic size which influence the retarding effect on electrons. These include ability to absorb electron energy on collision and trap electrons to form negative ions. The noble atomic gases (helium, argon, neon, etc.) are poorest in these respects and have the lowest dielectric strengths. Table 4-8 gives the relative dielectric strengths of a variety of gases at 1 atm pressure at a $p\,.\,d$ value of 1 atm × 0.25 cm. The relative strengths vary with the $p\,.\,d$ value, as well as gap geometry, and particularly in divergent fields where corona begins before breakdown. It is best to consult specific references with regard to divergent field breakdown values.

Corona and Breakdown in Nonuniform Fields between Conductors. In nonuniform fields, when the ratio of spacing to conductor radius of curvature is about 3 or less, breakdown occurs without prior corona. The breakdown voltage is controlled by the integral of the Townsend ionization coefficient α across the gap. At larger ratios of spacing to radius of curvature, corona discharge occurs at voltage levels below complete gap breakdown.

Corona in air at atmospheric pressure occurs before breakdown when the ratio of outer to inner radius of coaxial electrodes exceeds 2.72 or where the ratio of gap to sphere radius between spheres exceeds 2.04. These discharges project some distance from the small-radii conductor but do not continue out into the weaker electric field region until a higher voltage level is reached.

Such partial breakdowns are often characterized by rapid pulses of current and radio noise. With some conductors at intermediate voltages between onset and complete breakdown, they blend into a pulseless glow discharge around the conductor. When corona occurs before breakdown, it creates an ion space charge around the conductor, which modifies the electric field, reducing the stress at sharp conductor points in the intermediate voltage range. At higher voltages, streamers break out of the space-charge region and cross the gap.

The surface voltage stress at which corona begins increases above that for uniform field breakdown stress, since the field to initiate breakdown must extend over a finite distance. An empirical relation developed by Peek is useful for expressing the maximum surface stress for corona onset in air for several geometries of radius r cm:

$$\text{For concentric cylinders: } E = 31\delta\left(1 + \frac{0.308}{\sqrt{\delta r}}\right) \quad \text{kV/cm} \tag{4-74}$$

$$\text{For parallel wires: } E = 29.8\delta\left(1 + \frac{0.301}{\sqrt{\delta r}}\right) \quad \text{kV/cm} \tag{4-75}$$

$$\text{For spheres: } E = 27.2\delta\left(1 + \frac{0.54}{\sqrt{\delta r}}\right) \quad \text{kV/cm} \tag{4-76}$$

where δ is the density of air relative to that at 25°C and 1 atm pressure.

Corona Discharges on Insulator Surfaces. It has been shown by a number of investigators that the discharge-threshold voltage stress on or between insulator surfaces is the same as between metal electrodes. Thus, the threshold voltage for such discharges can be calculated from the series dielectric-capacitance relation for internal gaps of simple shapes, such as plane and coaxial gaps, insulated conductor surfaces, and hollow spherical cavities.

The corona-initiating voltage at a conductor edge on a solid barrier depends on the electric stress concentration and generally on the ratio of the barrier thickness to its dielectric constant, except with low surface resistance. Any absorbed water or conducting film raises the corona threshold voltage by reducing stress concentration at the conductor edge on the surface.

It is sometimes possible to overvolt such gaps considerably prior to the first discharge, and the offset voltage may be below the proper voltage due to surface-charge concentration. With ac voltages, pulse discharges occur regularly back and forth each half cycle, but with dc voltage, the first discharge deposits a surface charge on the insulator surface which must leak away before another discharge can occur. Thus, corona on or between insulator surfaces is very intermittent with steady dc voltages, but discharges occur when the voltage is raised or lowered.

Flashover on Solid Surfaces in Gases. As has been mentioned in the previous section on partial discharges, the breakdown in gases is influenced by the presence of solid insulation between conductors. This insulation increases the electric stress in the gas. A particular case of this is the complete breakdown between conductors across or around solid insulator surfaces. This can occur when the conductors are on the same side of the insulation or on opposite sides. A significant reduction in flashover voltage can occur whenever a significant part of the electric field passes through the insulation. The reduction is influenced by the percentage of electric flux which passes through the solid insulation and the dielectric constant of the insulation.

4.3.3 Insulating Oils and Liquids

General Considerations. Typical insulating liquids are natural or synthetic organic compounds and frequently consist of mixtures of essentially isomeric compounds with some range of molecular weight. The mixture of very similar but not exactly the same molecules, with a range of molecular size and with chain and branched hydrocarbons, prevents crystallization and results in a low freezing point, together with a relatively high boiling point. Typical insulating liquids have permittivities (dielectric constants) of 2 to 7 and a wide range of conductivities depending on their purity. The dc conductivity in these liquids is usually due to dissolved impurities, which are ionized by dissociation. Higher ionized impurity and conductivity levels occur in liquids having higher permittivities and lower viscosities.

The function of insulating liquids is to provide electrical insulation and heat transfer. As insulation, the liquid is used to displace air in the system and provide a medium of high electric strength to fill pores, cracks, and gaps in insulation systems. It is usually necessary to fill and impregnate systems with liquid under vacuum so that all air bubbles are eliminated. If air is completely displaced in all high-electric-field regions, the corona threshold voltage and breakdown voltage for the system are greatly increased. The viscosity selected for a liquid insulation is often a compromise to provide the best balance between electrical insulation and heat transfer and other limitations such as flammability, solidification at low temperatures, and pressure development at high temperatures in sealed systems.

The most commonly used insulating liquids are natural hydrocarbon mineral oils refined to give low conductivity and selected viscosity and vapor-pressure levels for transformer, circuit-breaker, and cable applications. A number of synthetic fluids are also used for particular applications where the higher cost above that of mineral oil is warranted by the requirements of the application or by the improved performance in relation to the apparatus design.

Mineral Insulating Oils. Mineral insulating oils are hydrocarbons (compounds of hydrogen and carbon) refined from crude petroleum deposits from the ground. They consist partly of aliphatic compounds with the general formula $C_nH_{2n} + 2$ and C_nH_2n, comprising a mixture of straight- and branched-chain and cyclic or partially cyclic compounds. Many oils also contain a sizable fraction of aromatic compounds related to benzene, naphthalene, and derivatives of these with aliphatic side chains. The ratio of aromatic to aliphatic components depends on the source of the oil and its refining treatment. The percentage of aromatics is of importance to the gas-absorption or evaluation characteristics under electrical discharges and to the oxidation characteristics.

The important physical properties of a mineral oil (as for other insulating liquids as well) are listed in Table 4-9 for three types of mineral oils. In addition to these properties, mineral oils which are exposed to air in their application have distinctive oxidation characteristics which vary with type of oil and additives and associated materials.

Many manufacturers now approve the use of any of several brands of mineral insulating oil in their apparatus provided that they meet their specifications which are similar to ASTM D1040, (values from which are tabulated in Table 4-9). Low values of dielectric strength may indicate water or dirt contamination. A high neutralization number will indicate acidity, developed very possibly from oxidation, particularly if the oil has used been already. Presence of sulfur is likely to lead to corrosion of metals in the oil.

The solubility of gases and water in mineral oil is of importance in regard to its function in apparatus. Solubility is proportional to the partial pressure of the gas above the oil

$$S = S_0(p/p_0) \tag{4-77}$$

TABLE 4-9 Characteristic Properties of Insulating Liquids

Type of liquid	Mineral oil		
	Transformer	Cable and capacitor	Solid cable
Specific gravity	0.88	0.885	0.93
Viscosity, Saybolt sec at 37.8°C	57–59	0.100	100
Flash point, °C	135	165	235
Fire point, °C	148	185	280
Pour point, °C	–45	–45	–5
Specific heat	0.425	0.412	....
Coefficient of expansion	0.00070	....	0.00075
Thermal conductivity, cal/(cm) (s) (°C)	0.39	....	....
Dielectric strength,* kV~	30	....	....
Permittivity at 25°C	2.2	....	....
Resistivity, $\Omega \cdot cm \times 10^{12}$	1–10	50–100	1–10

*ASTM D877.

where S is the amount dissolved at pressure p if the solubility is expressed as the amount S_0 dissolved at pressure p_0.

The solubility is frequently expressed in volume percent of the oil. Values for solubility of some common gases in transformer oil at atmospheric pressure (760 torr) and 25°C are air 10.8%, nitrogen 9.0%, oxygen 14.5%, carbon dioxide 99.0%, hydrogen 7%, and methane 30% by volume. The solubilities of all the gases, except CO_2, increase slightly with increasing temperature. Water is dissolved in new transformer oil to the extent of about 60 to 80 ppm at 100% relative humidity and 25°C. The amount dissolved is proportional to the relative humidity. Solubility of water increases with oxidation of the oil and the addition of polar impurities, with which the water becomes associated. Larger quantities of water can be suspended in the oil as fine droplets.

Dielectric Properties of Mineral Oils. The permittivity of mineral insulating oils is low, since they are essentially nonpolar, containing only a few molecules with electric dipole moments. Some oils possess a minor fraction of polar constituents, which have not been identified. These contribute a dipolar character to the dielectric properties at low temperature and/or high frequency. A typical permittivity for American transformer oil at 60 Hz is 2.19 at 25°C, declining almost linearly to 2.11 at 100°C. At low temperatures and high frequencies, values of permittivity as high as 2.85 have been noted in oils with a relatively high level of polar constituents.

The dc conductivity levels of mineral oils range from about $10^{-15}\ \Omega^{-1} \cdot \text{cm}^{-1}$ for pure new oils up to $10^{-12}\ \Omega^{-1} \cdot \text{cm}^{-1}$ for contaminated used oils. This conductivity is due to dissociated impurity ions or ions developed by oil oxidation. It increases approximately exponentially with temperature about 1 decade in 80°C.

Alternating-current dissipation-factor values are nearly proportional to the dc conductivity $10^{-13}\ \Omega^{-1}$, corresponding to a tan δ of 0.008. If no electrode polarization or interfacial polarization effects at solid barrier surface are present, the dc conductivity σ should be related to the ac conductivity (tan δ) by $\sigma = \frac{5}{9}\epsilon' f \tan\delta \times 10^{-12}$ where ϵ' is the dielectric permittivity (Table 4-7) and f is the frequency.

Corona or partial breakdown can occur in mineral oil, as with any liquid or gas, when the electric stress is locally very high and complete breakdown is limited by a solid barrier or large oil gap (as with a needle point in a large gap). Such discharges produce hydrogen and methane gas, and sometimes carbon with larger discharges. Dissolved air is also sometimes released by the discharge. If the gas bubbles formed are not ejected away from the high field, they will reduce the subsequent discharge threshold voltage to as much as 80%. The resistance of insulating oils to partial discharges is measured by two ASTM gassing tests: D2298 (Merrill test) and D2300 (modified Pirelli test). These tests measure the amount of decomposition gas evolved under specified conditions of exposure to partial discharges. A minimum amount of gas is, of course, preferred, particularly in applications for cables or capacitors. In fact, conventional mineral oils are inadequate in this respect for application in modern 60-Hz power capacitor designs.

Deterioration of Oil. Deterioration of oil in apparatus partially open or "breathing" is subject to air oxidation. This leads to acidity and sludge. There is no correlation between the amount of acid and the likelihood of sludging or the amount of sludge. Sludge clogs the ducts, reduces the heat transfer, and accelerates the rate of deterioration. ASTM tests for oxidation of oils are D1904, D1934, D1313, and D1314. Copper and lead and certain other metals accelerate the oxidation of mineral oils. Oils are considerably more stable in nitrogen atmospheres.

Inhibitors are now commonly added both to new and to used oils to delay the oxidation. Ditertiary butyl paracresol (DBPC) is the inhibitor most commonly used at present.

Servicing, Filtering, and Treating. Oil in service is usually maintained by testing for acidity, dielectric strength, inhibitor content, interfacial tension, neutralization number, peroxide number, pour point, power factor, refractive index and specific optical dispersion, resistivity, saponification, sludge, corrosive sulfur, viscosity, and water content, as outlined in ASTM D117. These properties indicate various types of contamination or deterioration which might affect the operation of the insulating oil.

Depending on the voltage rating of the apparatus, the oil is maintained above 16 to 22 kV (ASTM test D877). The usual contaminants are water, sludge, acids, and in circuit-breaker oils, carbon. The centrifuge is best suited for removing large quantities of water, heavier solid particles, etc. The blotter filter press is used for the removal of minute quantities of water, fine carbon, etc. In another method, after removing the larger particles, the oil is heated and sprayed into a vacuum chamber, where the water and volatile acids are removed. Sludge and very fine solids are then taken out by a blotter filter press. All units are assembled together so that the process is completed in a single pass. Some work has been done in reclaiming oil by treating it to reduce acidity. One process is similar to the later stages in refining. Another treatment uses activated alumina, Fuller's earth, or silica gel. The IEEE Guide for Maintenance of Insulating Oil is published as *IEEE Standards Publ*. 64.

It has been found that analysis of the dissolved gas in oil or above the oil in oil-insulated transformers and cables is a good diagnostic tool to detect electrical faults, particularly, or deterioration, generally. For example, continuing or intermittent partial discharges produce hydrogen and low-molecular-weight hydrocarbons such as methane, ethane, and ethylene which accumulate in the oil and can be measured accurately to assess the magnitude of the fault. Higher-current arc faults produce acetylene in addition to H_2 and other low-molecular-weight hydrocarbons. Thermal deterioration of cellulosic or paper insulation is indicated by elevated concentrations of CO and CO_2 in the oil.

Synthetic Liquid Insulation. Synthetic chlorinated diphenyl and chlorinated benzene liquids (askarels) have been used widely from the mid-1930s up to the mid-1970s and are still in service in many power capacitors and transformers, where they were adopted for their nonflammability as well as good electrical characteristics. Since the mid-1970s, their use has been banned in most countries due to their alleged toxicity and resistance to biodegradation in the environment. Now, when apparatus containing these fluids, which are commonly referred to as *PCBs*, are taken out of service, environmental regulations in the United States require that the fluid not be released into the environment. Waste fluid should be incinerated at high temperature with HCl reactive absorbent scrubbers in the stack, since this acid gas is a product of the combustion.

New synthetic fluids have been developed and are now widely applied in power capacitors where the electrical stresses are very high. These fluids include aromatic (containing benzene rings) hydrocarbons, some of which have excellent resistance to partial discharges. They are not fire-resistant, however.

Very high boiling, low-vapor-pressure, high-flash-point (>300°C) hydrocarbon oils are being tried for power transformers with some fire resistance. Methods for assessing the risk of fire with such liquids, as well as with silicones, are still being debated.

Perchlorethylene (tetrachloroethylene), a nonpolar liquid, is now in use in sealed medium-power transformers, where nonflammability is required. With a boiling point at atmospheric pressure of 121°C, this fluid is completely nonflammable. It is also widely used in dry cleaning. Other important classes of synthetic insulating fluids are discussed in the following sections.

Fluorocarbon Liquids. A number of nonpolar nonflammable perfluorinated aliphatic compounds, in which the hydrogen has been completely replaced by fluorine, are available with different ranges of viscosity and boiling point from below room temperature to more than 200°C. These compounds have low permittivities (near 2.0) and very low conductivity. They are inert chemically and have low solubilities for most other materials. The chemical formula for these compounds is one of the following: C_nF_{2n}, C_nF_{2n+2}, or $C_nF_{2n}O$. The presence of the oxygen in the latter formula does not seem to reduce the stability. These compounds have been used for filling electronic apparatus and large transformers to give high heat-transfer rates together with high dielectric strength. The vapors of these liquids also have high dielectric strengths.

Silicone Fluids. These fluids, chemically formed from Si—O chains with organic (usually methyl) side groups, have a high thermal stability, low temperature coefficient of viscosity, low dielectric losses, and high dielectric strength. They can be obtained with various levels of viscosity and correlated vapor pressures. Rated service temperatures extend from –65 to 200°C, some having short-time capability up to 300°C. Their permittivity is about 2.6 to 2.7, declining with increasing temperature. These fluids have a tendency to form heavier carbon tracks than other insulating liquids when breakdown occurs. They cannot be considered fireproof but will reduce the risk of fire due to their low vapor pressure.

Ester Fluids. There are a few applications, mostly for capacitors, where organic ester compounds are used. These liquids have a somewhat higher permittivity, in the range of about 4 to 7, depending on the ratio of ester groups to hydrocarbon chain lengths. Their conductivities are generally somewhat higher than those of the other insulating liquids discussed here. The compounds are easily subject to hydrolysis with water to form acids and alcohols and should be kept dry, particularly if the temperature is raised. Their thermal stability is poor. Specifically, dibutyl sebacate has been used in high-frequency capacitors and castor oil in energy-storage capacitors.

4.3.4 Insulated Conductors

Insulated conductors vary from those carrying only a few volts to those carrying thousands of volts. They range from low-voltage bell wire with conductor gage of 22 to 24 to power cables with conductors of 2000 kcmil or 1013 mm^2 in cross-sectional area. The conductors can be round, rectangular, braided, or stranded. They can be of aluminum or copper. The insulation can be thin as in magnet wire or thick as in underground or marine cables. The insulation system can vary with functional application. It can be extruded or taped. It can be thermoplastic or thermoset. It can be a polymer in combination with cotton or glass cloth. There can be several different layers with different functional roles. Some of the applications for insulated conductors are communications, control, bell, building, hookup, fixture, appliance, and motor lead. The insulation technology for magnet wire and for power cables has been studied extensively because of the severe stresses seen by these insulation systems.

Flexible Cords. Flexible cords and cables cover appliance and lamp cords, extension cords for home or industrial use, elevator traveling cables, decorative-lighting wires and cords, mobile home wiring, and wiring for appliances that get hot (e.g., hot plates, irons, cooking appliances). The requirements for these cables vary a great deal with application. They must be engineered to be water-resistant, impact-resistant, temperature-tolerant, flex-tolerant, linearly strong, and flame-resistant and have good electrical insulation characteristics.

Magnet Wire Insulation. The term *magnet wire* includes an extremely broad range of sizes of both round and rectangular conductors used in electrical apparatus. Common round-wire sizes for copper are AWG No. 42 (0.0025 in) to AWG No. 8 (0.1285 in). A significant volume of aluminum magnet wire is produced in the size range of AWG No. 4 to AWG No. 26. Ultrafine sizes of round wire, used in very small devices, range as low as AWG No. 60 for copper and AWG No. 52 for aluminum.

Approximately 20 different “enamels” are used commercially at present in insulating magnet wire. Magnet wire insulations are high in electrical, physical, and thermal performance and best in space factor. The most widely used polymers for film-insulated magnet wire are based on polyvinyl acetals, polyesters, polyamideimides, polyimides, polyamides, and polyurethanes. Many magnet wire constructions use different layers of these polymer types to achieve the best combination of properties. The most commonly used magnet wire is NEMA MW-35C, Class 200, which is constructed with a polyester basecoat and a polyamideimide topcoat. Polyurethanes are employed where ease of solderability without solvent or mechanical striping is required. The thermal class of polyurethane insulations has been increased up to Class 155 and even Class 180. Magnet wire products also are produced with fabric layers (fiberglass or Dacron-fiberglass) served over bare or conventional film-insulated magnet wire. Self-bonding magnet wire is produced with a thermoplastic cement as the outer layer, which can be heat-activated to bond the wires together.

Power Cables. Insulated power cables are used extensively in underground residential distribution. There has been extensive replacement of PILC, or paper in lead cable, with extruded polymer-insulated cables. Although PILC is still dominant for underground transmission cables, extruded polymeric cables are also beginning to be used for these high-voltage applications. Typical cable sizes with the cross section of the conductor are shown in the Table 4-10.

TABLE 4-10 Typical Cable Sizes

Cable size	Conductor cross section, mm^2
AWG 2	33.6
AWG 1	42.4
AWG 1/0	53.5
AWG 2/0	67.4
AWG 3/0	85.0
AWG 4/0	107.2
500 kcmil	253.5
750 kcmil	379.5
1000 kcmil	507.0
2000 kcmil	1013.0

Typically, a cable rated at 15 kV will have insulation of wall thickness 175 mil (4.45 mm); one rated at 35 kV will have a wall thickness of 345 mil (8.76 mm); one rated at 69 kV will have insulation thickness of 650 mil (16.5 mm); and a 138-kV cable will have insulation of wall thickness 850 mil (21.6 mm). A cable construction includes the conductor shield, insulation, and insulation shield. In addition, most cables these days have a jacket to diminish moisture penetration into the insulation.

The conductor shield is a semiconductive material applied to the conductor to smooth out the stress. Since the conductors, especially the stranded conductors, have "bumps" that can enhance the field, the role of the semiconductor is to present an even voltage stress to the insulation. The insulation shield fulfills a similar role on the outer surface of the insulation. Grit, or especially metal particles, can be sites where breakdown begins. A clean interface and a semiconductive material prevent such sites from forming. The formulation of the conductor shield and the insulation shield is different. The formulation also depends on the insulating material used. A number of different materials have been used as the matrix material for semiconductive shields. These include low-density polyethylene (LDPE), ethylene–ethyl acrylate (EEA), ethylene–vinyl acetate (EVA), ethylene–propylene rubber (EPR), ethylene–propylene diene monomer (EPDM), butyl rubber, and various proprietary formulations. These materials, in themselves, are not conducting. They are made conducting by loading the polymer with carbon.

There are two insulations in use for power cables. One is cross-linked polyethylene (XLPE) and the other is ethylene–propylene rubber (EPR). These insulating materials will be described in greater detail in the following paragraphs.

Most of the cables being installed in the latter part of the 1990s are jacketed. The jacket provides protection against oil, grease, and chemicals. However, the primary role played by the jacket is to slow down the ingress of moisture, since moisture in the presence of an electric field causes the insulation to degrade by a process called *treeing*. One of the materials used extensively as a jacket material is linear low-density polyethylene (LLDPE). Jackets are approximately 50 mil (1.27 mm) thick. The discussion thus far has not described the chemistry of each of these insulating materials. The terms *thermoset* and *thermoplastic* are used without explanation. Material names such as PE, PTFE, PVC, and silicones are used without characterizing the chemistry or structure.

A *thermoplastic* resin is one with a melting point. With rising temperature, a thermoplastic resin first undergoes a glass-transition temperature (T_g) and then a crystalline melting point (T_m). Below the glass-transition temperature, a polymer is rigid and exhibits properties associated with the crystalline state. Above the glass-transition temperature, the material becomes plastic and viscous, and the material starts to slowly approach the structure of the liquid state. The glass-transition state can be detected by plotting the dielectric constant, refractive index, specific heat, coefficient of expansion, or electrical conductivity as a function of temperature.

There is one characteristic slope below the glass-transition state and another steeper slope above the glass-transition temperature. Approximate values for T_g and T_m for polyethylene are –128 and 115°C, and for polystyrene they are 80 and 240°C. It is difficult to give exact values for a given generic polymer. This is so because the exact value will depend a great deal on the variation in the character of a particular polymer, with all the variations being grouped together and called by a common name. For example, for polyethylene, the molecular weight (the degree of polymerization) of the resin, the degree of branching, and the size or length of the branches will affect both T_g and T_m. With polyvinyl chloride, the steroregularity, copolymerization, and plasticization all will affect T_g and T_m.

A *thermoset* resin does not exhibit a visible melting point. An epoxy or a phenolic resin has a three-dimensional network structure. The three-dimensional structure results in a rigid framework that cannot be made fluid without breaking a large number of bonds. The thermoplastic resins, on the other hand, are linear. They might be thought of as strands of spaghetti. The strands can slip by one another and can be fluid. The analogy to a bowl of spaghetti can be used in understanding how a thermoset material can be formed by cross-linking a thermoplastic polymer such as polyethylene. The cross-linking reaction forms bonds between the linear strands of polyethylene to form a three-dimensional structure. To visualize the cross-linking, an analogy that can be used is that of the bowl of spaghetti left in a refrigerator overnight. Once the strands of spaghetti stick together, the mass is no longer fluid. The mass can be taken out of the bowl, and it will retain the shape of the bowl. The only way to fluidize this spaghetti is to break most or all the bonds formed between the individual strands.

Some of the insulation materials used for insulated conductors are polyethylene (PE), ethylene–propylene rubber (EPR), polyvinyl chloride (PVC), fluorinated ethylene propylene (FEP), ethylene chlorotrifluoroethylene, polytetrafluoroethylene (PTFE), butyl rubber, neoprene, nitrile–butadiene rubber (NBR), latex, polyamide, and polyimide.

Polyethylene is made by polymerizing ethylene, a gas with a boiling point of –104°C. A reaction carried out at high temperature (up to 250°C) and high pressure (between 1000 and 3000 atm) produces low-density polyethylene. The reason for the low density is that the short and long branches on the long chains prevent the chains from packing efficiently into a crystalline mass. The use of Ziegler-Natta catalysts results in high-density polyethylene. The use of the catalyst results in less branching and thereby a polymer that can pack more efficiently into crystalline domains. Recently, shape-selective catalysts have become available that produce polyethylene polymers that can be made with designer properties. Even though polyethylene consists of chains of carbons, the properties can vary depending on molecular weight and molecular shape. Polyethylene sold for insulating purposes has only small amounts of additives. There is always some antioxidant. For cross-linked polyethylene, the residues of the cross-linking agent are present. Additives to inhibit treeing are added.

Ethylene–propylene rubber is a copolymer made from ethylene and propylene. The physical properties of the neat polymer are such that it is not useful unless compounded. The finished compounded product has as much as 40 to 50% filler content. Fillers consist of clays, calcium carbonate, barium sulfate, or various types of silica. In addition to the filler, EPR is compounded with plasticizer, antioxidants, flame retardants, process aids, ion scavengers, coupling agents, a curing coagent, and a curative.

Polyvinyl chloride is a polymer made from vinyl chloride, a gas boiling at –14°C. It is partially syndiotactic; that is, the stereochemistry of the carbons on which the chlorines are attached is more or less alternating. By being only partially syndiotactic, the crystallinity is low. However, the polymer is still fairly rigid, and for use where flexibility is desired, the polymer must be plasticized.

Dibutylphthalate is often used as a plasticizer. In addition to plasticizers, PVC contains heat and light stabilizers. Oxides, hydroxides, or fatty acid salts of lead, barium, tin, or cadmium are typical stabilizers.

Polytetrafluoroethylene (or Teflon) is a polymer made from tetrafluoroethylene, a nontoxic gas boiling at –76°C. It is a linear polymer consisting of chains made of CF_2 units. Its crystallinity is quite high, and its crystalline melting point is 327°C. It is resistant to almost all reagents, even up to the boiling point of the reagent. It is attacked only by molten alkali metals or the alkali metal dissolved in liquid ammonia. Polytetrafluoroethylene exhibits excellent electrical properties. It has a low dielectric constant and a low loss factor. These electrical properties do not change even when the polymer is kept at 250°C for long periods of time.

Fluorinated ethylene propylene is a copolymer made from tetrafluoroethylene and hexafluoropropylene. It compares in toughness, chemical inertness, and heat stability to polytetrafluoroethylene (PTFE).

Polychlorotrifluoroethylene has performance properties that are surpassed only by PTFE and FEP. The crystalline melting point is 218°C, as compared with 327°C for PTFE. It retains useful properties to 150°C, as opposed to 250°C for PTFE. The advantage for polychlorotrifluoroethylene is that its melt viscosity is so low enough that molding and extrusion become more feasible than for PTFE and FEP.

Polyamides or Nylons are long-chain linear polymers made by molecules linked by amide linkages. Nylon 66 is made from hexamethylene diamine and adipic acid. Nylon 66 exhibits high strength, elasticity, toughness, and abrasion resistance. Nylon 6 is made from caprolactam, a cyclic amide. To form a polymer, the caprolactam opens and the amine group and carboxylic acid group form intermolecular amide links rather than the intramolecular amide link in the cyclic compound.

Polyimides are polymers connected by imide bonds. An amide is formed when the OH group of a carboxylic acid is replaced by the NH of an amine. An imide is a related structure formed when the noncarbonyl oxygen of an acid anhydride is replaced by a nitrogen of an amine. A polyimide is usually formed from an aromatic diamine and an aromatic dianhydride. The aromatic nature of the polyimide imparts thermal stability.

Rubbers used for electrical insulation can be either natural rubber or one of the synthetic rubbers. Natural rubber is obtained from the latex of different plants. The primary commercial source is the tree *Hevea brasiliensis*. Natural rubber is an isoprenoid compound wherein the isoprene (2-methyl-1,3-butadiene) is the unit of a high-molecular-weight polymer with a degree of polymerization of around 5000. Rubber without processing is too gummy to be of practical use. It is vulcanized (cross-linked) by reaction with sulfur. Natural rubber is flexible and elastic and exhibits good electrical characteristics.

Butyl rubbers are synthetic rubbers made by copolymerizing isobutylene (2-methyl-1-propene) with a small amount of isoprene. The purpose of isoprene is to introduce a double bond into the polymer chain so that it can be cross-linked. Butyl rubbers are mostly amorphous, with crystallization taking place on stretching. They are characterized by showing a low permeability to gases, thus making them the material of choice for inner tubes of automobile tires. They are reasonably resistant to oxidative aging. Butyl rubbers have good electrical properties.

Polychloroprene or *neoprene* is a generic term for polymers or copolymers of chloroprene (2-chloro-1,3-butadiene). Neoprene is an excellent rubber with good oil resistance. It has resistance to oxidative degradation, and is stable at high temperatures. Its properties are such that it would make excellent automobile tires, but the cost of the polymer makes it noncompetitive for this market. Its desirable properties are exploited for wire and cable insulations.

Nitrile rubbers are polymers of butadiene and acrylonitrile. Nitrile rubbers are used where oil resistance is needed. The degree of oil resistance varies with acrylonitrile content of the copolymer. With 18% acrylonitrile content, the oil resistance is only fair. With 40% acrylonitrile content, the oil resistance is excellent. The oil resistance is characterized by retention of low swelling, good tensile strength, and good abrasion resistance after being immersed in gasoline or oil. Nitrile rubbers can be used in contact with water or antifreeze. For use in wire insulation where oil resistance is needed, nitrile rubber is slightly better than neoprene.

4.3.5 Thermal Conductivity of Electrical Insulating Materials

One of the general characteristics of electrical insulating materials is that they are also good thermal insulating materials. This is true, in varying degrees, for the entire spectrum of insulating materials, including air, fluids, plastics, glasses, and ceramics. While the thermal insulating properties of electrical insulating materials are not especially important for electrical and electronic designs which are not heat sensitive, modern designs are increasingly heat sensitive. This is often because higher power levels are being dissipated from smaller part volumes, thus tending to raise the temperature of critical elements of the product design. This results in several adverse effects, including degradation of electrical performance and degradation of many insulating materials, especially insulating papers and plastics. The net result is reduced life and/or reduced reliability of the electrical or electronic part. To maximize life and reliability, much effort has been devoted to data and guidelines for gaining the highest possible thermal conductivity, consistent with optimization of product design limitations such as fabrication, cost, and environmental stresses. This section will present data and guidelines which will be useful to electrical and electronic designers in selection of electrical insulating materials for best meeting thermal design requirements. Also, methods of determining thermal conductivity K will be described.

Basic Thermal-Conductivity Data. The thermal-conductivity values for a range of materials commonly used in electrical design are shown in Table 4-11. These data show the ranking of the range of materials, both conductors and insulating materials, from high to low. The magnitude of the differences in conductor and plastic thermal-conductivity values can be seen. Note that one ceramic, 95% beryllia, has a higher thermal-conductivity value than some metals—thus making beryllia highly considered for high-heat-dissipating designs which allow its use. Thermal conductivity is variously reported in many different units, and convenient conversions are shown in Table 4-12.

Values of thermal conductivity do not change drastically up to 100°C or higher, and hence only a single value is usually given for plastics. For higher-temperature applications, such as with ceramics, the temperature effect should be considered. In addition to bulk insulating materials, insulating coatings are frequently used.

TABLE 4-11 Thermal Conductivity of Materials Commonly Used for Electrical Design

Material	Thermal conductivity	
	W/(in)(°C)	Btu/(h)(ft)(°F)
Silver	10.6	241
Copper	9.6	220
Eutectic bond	7.50	171.23
Gold	7.5	171
Aluminum	5.5	125
Beryllia 95%	3.9	90.0
Molybdenum	3.7	84
Cadmium	2.3	53
Nickel	2.29	52.02
Silicon	2.13	48.55
Palladium	1.79	40.46
Platinum	1.75	39.88
Chromium	1.75	39.88
Tin	1.63	36.99
Steel	1.22	27.85
Solder (60–40)	0.91	20.78
Lead	0.83	18.9
Alumina 95%	0.66	15.0
Kovar	0.49	11.1
Epoxy resin, BeO-filled	0.088	2.00
Silicone RTV, BeO-filled	0.066	1.5
Quartz	0.05	1.41
Silicon dioxide	0.035	0.799
Borosilicate glass	0.026	0.59
Glass frit	0.024	0.569
Conductive epoxy	0.020	0.457
Sylgard resin	0.009	0.21
Epoxy glass laminate	0.007	0.17
Doryl cement	0.007	0.17
Epoxy resin, unfilled	0.004	0.10
Silicone RTV, BeO-filled	0.004	0.10
Air		0.016

TABLE 4-12 Thermal-Conductivity Conversion Factors

From	To: $\frac{(\text{cal})(\text{cm})}{(\text{s})(\text{cm}^2)(°\text{C})}$	$\frac{(\text{W})(\text{cm})}{(\text{cm}^2)(°\text{C})}$	$\frac{(\text{W})(\text{in})}{(\text{in}^2)(°\text{C})}$	$\frac{(\text{Btu})(\text{ft})}{(\text{h})(\text{ft}^2)(°\text{F})}$
$\frac{(\text{cal})(\text{cm})}{(\text{s})(\text{cm}^2)(°\text{C})}$	1	4.18	10.62	241.9
$\frac{(\text{W})(\text{cm})}{(\text{cm}^2)(°\text{C})}$	2.39×10^{-1}	1	2.54	57.8
$\frac{(\text{W})(\text{in})}{(\text{in}^2)(°\text{C})}$	9.43×10^{-2}	3.93×10^{-1}	1	22.83
$\frac{(\text{Btu})(\text{ft})}{(\text{h})(\text{ft}^2)(°\text{F})}$	4.13×10^{-3}	1.73×10^{-2}	4.38×10^{-2}	1

Thermal-Conductivity Measurements. The recognized primary technique for measuring thermal conductivity of insulating materials is the guarded-hot-plate method (ASTM C177). A schematic of the apparatus is shown in Fig. 4-21. The purpose of the guard heater is to prevent heat flow in all but the axial (up and down in the schematic) direction by establishing isothermal surfaces on the specimen's hot side. With this condition established and by measuring the temperature difference across the sample, the electrical power to the main heater area and the sample thickness, the K factor, can be calculated as

$$K = \frac{QX}{2A\,\Delta T} \tag{4-78}$$

Instruments are available for this test which use automatic means to control the guard temperature and record the sample ΔT. Unfortunately this test is fairly expensive.

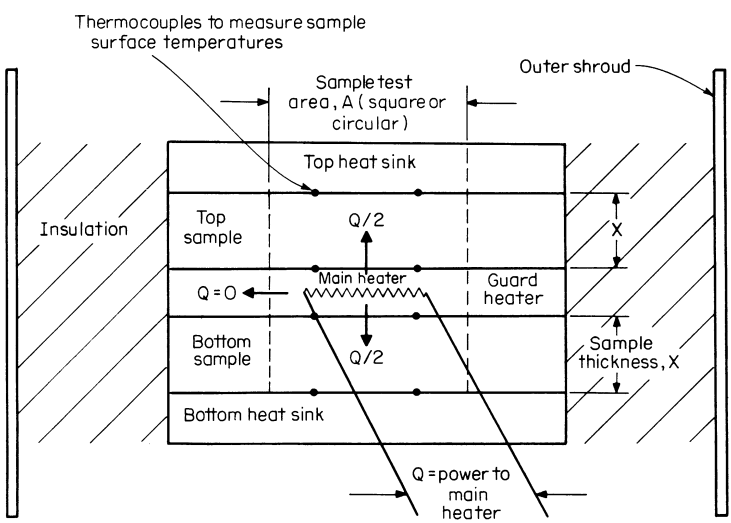


FIGURE 4-21 Schematic assembly of guarded hot plate.

Another technique uses a heat-flow sensor, which is a calibrated thermopile, in series with the heater, specimen, and cold sink. This method avoids the guard heater and requires only one specimen. This secondary technique is described in ASTM C518.

4.4 STRUCTURAL MATERIALS

4.4.1 Definitions of Properties

Stress. *Stress* is the intensity at a point in a body of the internal forces or components of force that act on a given plane through the point. Stress is expressed in force per unit of area (pounds per square inch, kilograms per square millimeter, etc.). There are three kinds of stress: tensile, compressive, and shearing. *Flexure* involves a combination of tensile and compressive stress. *Torsion* involves shearing stress. It is customary to compute stress on the basis of the original dimensions of the cross section of the body, though "true stress" in tension or compression is sometimes calculated from the area of the time a given stress exists rather than from the original area.

Strain. *Strain* is a measure of the change, due to a force, in the size or shape of a body referred to its original size or shape. Strain is a nondimensional quantity but is frequently expressed in inches per inch, etc. Under tensile or compressive stress, strain is measured along the dimension under consideration. *Shear strain* is defined as the tangent of the angular change between two lines originally perpendicular to each other.

Stress-Strain Diagram. A *stress-strain diagram* is a diagram plotted with values of stress as ordinates and values of strain as abscissas. Diagrams plotted with values of applied load, moment, or torque as ordinates and with values of deformation, deflection, or angle of twist as abscissas are sometimes referred to as stress-strain diagrams but are more correctly called *load-deformation diagrams*. The stress-strain diagram for some materials is affected by the rate of application of the load, by cycles of previous loading, and again by the time during which the load is held constant at specified values; for precise testing, these conditions should be stated definitely in order that the complete significance of any particular diagram may be clearly understood.

Modulus of Elasticity. The *modulus of elasticity* is the ratio of stress to corresponding strain below the proportional limit. For many materials, the stress-strain diagram is approximately a straight line below a more or less well-defined stress known as the *proportional limit*. Since there are three kinds of stress, there are three moduli of elasticity for a material, that is, the modulus in tension, the modulus in compression, and the modulus in shear. The value in tension is practically the same, for most ductile metals, as the modulus in compression; the modulus in shear is only about 0.36 to 0.42 of the modulus in tension. The modulus is expressed in pounds per square inch (or kilograms per square millimeter) and measures the elastic *stiffness* (the ability to resist elastic deformation under stress) of the material.

Elastic Strength. To the user and the designer of machines or structures, one significant value to be determined is a *limiting stress below which the permanent distortion of the material is so small that the structural damage is negligible and above which it is not negligible*. The amount of plastic distortion which may be regarded as negligible varies widely for different materials and for different structural or machine parts. In connection with this limiting stress for elastic action, a number of technical terms are in use; some of them are

1. *Elastic Limit*. The greatest stress which a material is capable of withstanding without a permanent deformation remaining on release of stress. Determination of the elastic limit involves repeated application and release of a series of increasing loads until a set is observed upon release of load. Since the elastic limit of many materials is fairly close to the proportional limit, the latter is sometimes accepted as equivalent to the elastic limit for certain materials. There is, however, no fundamental relation between elastic limit and proportional limit. Obviously, the value of the elastic limit determined will be affected by the sensitivity of apparatus used.

2. *Proportional Limit*. The greatest stress which a material is capable of withstanding without a deviation from proportionality of stress to strain. The statement that the stresses are proportional to strains below the proportional limit is known as *Hooke's Law*. The numerical values of the proportional limit are influenced by methods and instruments used in testing and the scales used for plotting diagrams.
3. *Yield Point*. The lowest stress at which marked increase in strain of the material occurs *without increase in load*. If the stress-strain curve shows no abrupt or sudden yielding of this nature, then there is no yield point. Iron and low-carbon steels have yield points, but most metals do not, including iron and low-carbon steels immediately after they have been plastically deformed at ordinary temperatures.
4. *Yield Strength*. The stress at which a material exhibits a specified limiting permanent set. Its determination involves the selection of an amount of permanent set that is considered the maximum amount of plastic yielding which the material can exhibit, in the particular service condition for which the material is intended, without appreciable structural damage. A set of 0.2% has been used for several ductile metals, and values of yield strength for various metals are for 0.2% set unless otherwise stated. On the stress-strain diagram for the material (Fig.4-22) this arbitrary set is laid off as *q* along the strain axis, and the line *mn* drawn parallel to *OA*, the straight portion of the diagram. Since the stress-strain diagram for release of load is approximately parallel to *OA*, the intersection *r* may be regarded as determining the stress at the yield strength. The yield strength is generally used to determine the elastic strength for materials whose stress-strain curve in the region *pr* is a smooth curve of gradual curvature.

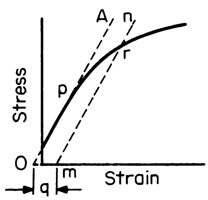


FIGURE 4-22 Yield strength of a material having no well-defined yield point.

Ultimate Strength. *Ultimate strength (tensile strength* or *compressive strength)* is the maximum stress which a material will sustain when slowly loaded to rupture. Ultimate strength is computed from the maximum load carried during a test and the original cross-sectional area of the specimen. For materials that fail in compression with a shattering fracture, the compressive strength has a definite value, but for materials that do not fracture, the compressive strength is an arbitrary value depending on the degree of distortion which is regarded as indicating complete failure of the material. In tensile tests of many materials, especially those having appreciable ductility, failure does not occur at the stress corresponding to the ultimate strength. For such materials, localized deformation, or necking, occurs and the nominal stress decreases because of the rapidly decreasing cross-sectional area until failure occurs.

Shearing Strength. *Shearing strength* is the maximum shearing stress which a material is capable of developing. The remarks in the preceeding paragraph regarding methods of failure are also applicable to failures in shear. Owing to experimental difficulties of obtaining true shearing strength, the values of modulus of rupture in torsion are usually reported as indicative of the shearing strength.

Modulus of Rupture. *Modulus of rupture* in flexure (or torsion) is the term applied to the computed stress, in the extreme fiber of a specimen tested to failure under flexure (or torsion), when computed by the arbitrary application of the formula for stress with disregard of the fact that the stresses exceed the proportional limit. Hence, the modulus of rupture does not give the true stress in the member but is useful only as a basis of comparison of relative strengths of materials.

Ductility. *Ductility* is that property of a material which enables it to acquire large permanent deformation and at the same time develop relatively large stresses (as drawing into a wire). Although ductility is a highly desirable property required by almost all specifications for metals, the quantitative amount needed for structural applications is not entirely clear but probably does not exceed about 3% elongation after the structure is fabricated. The commonly used measures of ductility are

1. *Elongation* is the ratio of the increase of length of a specimen, after rupture under tensile stress, to the original gage length; it is usually expressed in percent. The percentage of elongation for any given material depends upon the gage length, which should always be specified.
2. *Reduction of area or contraction of area* is the ratio of the difference between the original and the fractured cross section to the original cross-sectional area; it is usually expressed as a percentage.
3. *Bend test* measures the angle through which a given specimen of material can be bent, at a specified temperature, without cracking. In some cases, the maximum angle through which the specimen can be bent around a certain diameter or the number of bendings back and forth through a stated angle are measured. In other cases, the elongation in a given gage length across the crack on the tension side of the bend specimen is measured.

Plasticity. *Plasticity* permits a material to assume permanent deformations under loads without recovery of the strain when the loads are removed. Plasticity permits shaping of metal parts by plastic deformation; plastic materials deform instead of fracturing under load.

Brittleness. *Brittleness* is defined as the ability of a material to fracture under stress with little or no plastic deformation. Brittleness implies a lack of plasticity.

Resilience. *Resilience* is the amount of strain energy (or work) which may be recovered from a stressed body when the loads causing the stresses are removed. Within the elastic limit, the work done in deforming the bar is completely recovered upon removal of the loads; the total amount of work done in stressing a unit volume of the material to the elastic limit is called the *modulus of resilience*.

Toughness. *Toughness* is the ability to withstand large stresses accompanied by large strains before fracture. The toughness is usually measured by the total work done in stressing a unit volume of the material to complete fracture and may be interpreted as the total area under the stress-strain curve. Ductility differs from toughness in that it deals only with the ability of the material to deform, whereas toughness is measured by the energy-absorbing capacity of the material.

Impact Resistance. The ability of a material to resist impact or energy loads without permanent distortion is measured by the modulus of resilience. The ultimate resistance to impact before fracture is measured by the toughness of the material. For members with abrupt changes of section (holes, keyways, fillets, etc.), the resistance to a rapidly applied load depends greatly on the *notch sensitivity* (the resistance to the formation and spread of a crack); above certain critical velocities of loading and below certain critical temperatures, the impact strength is greatly reduced. Relative notch sensitivity under repeated loads is not the same as that in a single-blow notched-bar test. Impact values are influenced by speed of straining, shape and size of specimen, and type of testing machine.

Charpy or *Izod* impact bend tests measure the energy required to fracture small notched specimens (1 cm^2) under a single blow. These tests are used as an indication of toughness, a property that is very sensitive to the composition and thermal-mechanical history of the material. Tests should be carried out over a range of temperatures to determine the temperature at which the alloy fails by brittle rather than ductile failure.

Fracture Mechanics. Three primary factors have been identified that control the susceptibility of a structure to brittle failure: material toughness (affected by composition and metallurgical structure as well as temperature, strain rate, and constraints to plastic yielding), flaw size (internal discontinuities such as porosity or small cracks from welding, fatigue, and fabrication), and stress level (applied or residual). Fracture mechanics attempt to interrelate these variables in order to predict the occurrence of brittle fracture on a quantitative design basis rather than depend on qualitative relationships between experience and results of impact tests such as Izod and Charpy. Fracture mechanics has had excellent success when applied to high-strength materials. The material parameter defined, called the *fracture toughness* K_C, can be measured experimentally and used to specify safe loading conditions in the presence of a given size and geometry of flaw.

Hardness. *Hardness* is the resistance which a material offers to small, localized plastic deformations developed by specific operations such as scratching, abrasion, cutting, or penetration of the surface. Hardness does not imply brittleness, as a hard steel may be tough and ductile. The standard Brinell hardness test is made by pressing a hardened steel ball against a smooth, flat surface under certain standard conditions; the Brinell hardness number is the quotient of the applied load divided by the area of the surface of the impression. A different method of test is employed in the Shore scleroscope, in which a small, pointed hammer is allowed to fall from a definite height onto the material, and the hardness is measured by the height of the rebound, which is automatically indicated on a scale. The Rockwell hardness machine measures the depth of penetration in the metal produced by a definite load on a small indenter of spherical or conical shape. Vickers or Tukon hardness machines measure hardness on a microscopic scale. The dimensions of the impression of a lightly loaded diamond pyramid indenter on a polished surface are related to hardness number.

Fatigue Strength. *Fatigue strength (fatigue limit)* is a limiting stress below which no evidence of failure by progressive fracture can be detected after the completion of a very large number of repetitions of a definite cycle of stress. The fatigue limits usually reported are those for completely reversed cycles of flexural stress in polished specimens. For stress cycles in which an alternating stress is superimposed on a steady stress, the endurance limit (based on the maximum stress in the cycle) is somewhat higher. Most ferrous metals have well-defined limits, whereas the fatigue strength of many nonferrous metals is arbitrarily listed as the maximum stress that is just insufficient to cause fracture after some definite number of cycles of stress, which should always be stated. The fatigue strength of actual members containing notches (holes, fillets, surface scratches, etc.) is greatly reduced and depends entirely on the "stress-raising" effect of these discontinuities and the sensitivity of the material to the localized stresses at the notch.

Composition and Structure. Chemical analysis is employed to determine whether component elements are present within specified amounts and impurity elements are held below specified limits. Mechanical and physical properties, however, depend on the size, shape, composition, and distribution of the crystalline constituents that make up the structure of the alloy. Chemical analysis does not reveal these features of the structure. Metallographic techniques, which involve examination of carefully polished and etched surfaces by optical and electron microscopy or x-ray methods, are required to provide this vital information. Nondestructive testing (NDT) methods are useful in detecting the presence of flaws of various kinds in finished parts and structures. These techniques depend on the interference of the defect with some easily measured physical property, such as x-ray absorption, magnetic susceptibility, propagation of acoustical waves, or electrical conductivity. NDT techniques have particular application where defects are difficult to detect and quite likely to occur (as in welded structures), and where high-integrity performance requires 100% inspection.

Aging. *Aging* is a spontaneous change in properties of a metal with time after a heat-treatment or a cold-working operation. Aging tends to restore the material to an equilibrium condition and to remove the unstable condition induced by the prior operation, and usually results in increased strength of the metal with corresponding loss of ductility. The fundamental action involved is generally one of precipitation of hardening elements from the solid solution, and the process can usually be hastened by slight increase in temperature. This is a very important strengthening mechanism in a variety of ferrous and nonferrous alloys, for example, high-strength aluminum alloys.

Corrosion Resistance. There is no universal method of determining corrosion resistance, because different types of exposures ordinarily produce entirely dissimilar results on the same material. In general, the subject of corrosion is rather complicated; in some cases, corrosive attack appears to be chiefly chemical in nature, while in others the attack is by electrolysis. Owing to the great diversity of materials exposed to corrosive influences in service and the wide range of service conditions, it is impracticable to formulate any universal measure of corrosion resistance. If the service life is likely to be determined by corrosion resistance, the degree of impairment which marks the end of usefulness ordinarily will be established by considerations of safety and reliability or perhaps of appearance. Corrosion testing is conducted in general by two methods: (1) normal exposure in service with periodic observations of corrosive action as it progresses under such conditions, and (2) some type of artificially accelerated test, which may serve merely to obtain comparative results or, again, may simulate the conditions of service exposure.

Powder Metallurgy. Many alloys and metallic aggregates having unusual and very valuable properties are being produced commercially by mixing metal powders, pressing in dies to desired shapes, and sintering at high temperatures. Parts may be produced to close dimensional tolerance, and the process enables the mixing of dissimilar materials which will not normally alloy or which cannot be cast because of insolubility of the constituents. Wide use of powder metallurgy is made in producing copper-molybdenum alloys for contact electrodes for spot welding, extremely hard cemented tungsten carbide tips for use in metal cutting tools, and copper-base alloys containing either graphite particles or a controlled dispersion of porosity for bearings of the “oilless” or oil-retaining types. Silver-nickel and silver-molybdenum alloys (tungsten or graphite may be added) for contact materials having high conductivity but good resistance to fusing can be produced by the method. Powdered iron is being used to manufacture gears and small complex parts where the savings in weight of metal and machining costs are able to offset the additional cost of metal and processing in the powdered form. Small Alnico magnets of involved shape which are exceedingly difficult to cast or machine can be produced efficiently from metallic powders and require little or no finishing. Solid mixtures of metals and nonmetals, such as asbestos, can be produced to meet special requirements. The size and shape of powder particles, pressing temperature and pressure, sintering temperature and time, all affect the final density, structure, and physical properties.

4.4.2 Structural Iron and Steel

Classification of Ferrous Materials. Iron and steel may be classified on the basis of composition, use, shape, method of manufacture, etc. Some of the more important ferrous alloys are described in the sections below.

Ingot iron is commercially pure iron and contains a maximum of 0.15% total impurities. It is very soft and ductile and can undergo severe cold-forming operations. It has a wide variety of applications based on its formability. Its purity results in good corrosion resistance and electrical properties, and many applications are based on these features. The average tensile properties of Armco ingot iron plates are tensile strength 320 MPa (46,000 lb/in^2); yield point 220 MPa (32,000 lb/in^2); elongation in 8 in, 30%; Young’s modulus 200 GPa (29×10^6 lb/in^2).

Plain carbon steels are alloys of iron and carbon containing small amounts of manganese (up to 1.65%) and silicon (up to 0.50%) in addition to impurities of phosphorus and sulfur. Additions up to 0.30% copper may be made in order to improve corrosion resistance. The carbon content may range from 0.05% to 2%, although few alloys contain more than 1.0%, and the great bulk of steel tonnage contains from 0.08% to 0.20% and is used for structural applications. Medium-carbon steels contain around 0.40% carbon and are used for constructional purposes—tools, machine parts, etc. High-carbon steels have 0.75% carbon or more and may be used for wear and abrasion-resistance applications such as tools, dies, and rails. Strength and hardness increase in proportion to the carbon content while ductility decreases. Phosphorus has a significant hardening effect in low-carbon steels, while the other components have relatively minor effects within the limits they are found. It is difficult to generalize the properties of steels, however, since they can be greatly modified by cold working or heat treatment.

High-strength low-alloy steels are low-carbon steels (0.10% to 0.15%) to which alloying elements such as phosphorus, nickel, chromium, vanadium, and niobium have been added to obtain higher strength. This class of steel was developed primarily by the transportation industry to decrease vehicle weight, but the steels are widely used. Since thinner sections are used, corrosion resistance is more important, and copper is added for this purpose.

Free-Machining Steels. Additions of manganese, phosphorus, and sulfur greatly improve the ease with which low-carbon steels are machined. The phosphorus hardens the ferrite, and the manganese and sulfur combine to form nonmetallic inclusions that help form and break up machining chips. The improvement in machinability is gained at some loss of mechanical properties, and these steels should be used for noncritical applications. Small amounts of lead also improve machining characteristics of steel by helping break up chips as well as providing a self-lubricating effect. Lead is more often added to higher-carbon steels where the effect on mechanical properties is less detrimental than that caused by sulfide inclusions.

Alloy Steels. When alloying ingredients (in addition to carbon) are added to iron to improve its mechanical properties, the product is known as an *alloy steel*. Heat treatment is a necessary part of the manufacture and use of alloy steels; only through proper quenching and tempering can the full beneficial effects of the alloys be obtained. The chief advantages obtained from the addition of alloys to steel are (1) to increase the depth of hardening on quenching, thus making it possible to produce more uniform properties throughout thick sections with a minimum of distortion, and (2) to form chemical compounds which when properly distributed develop desirable properties in the steel, that is, extreme hardness, corrosion or heat resistance, and high strength without excessive brittleness.

The most commonly used alloy steels have been classified by the American Iron and Steel Institute and the Society of Automotive Engineers and are identified by a nomenclature system that is partially descriptive of the composition. The system of steel designations and the approximate strengths of several alloy steels after specific heat treatments are available from various manufacturers. Mechanical properties of the alloy steels vary over a wide range depending on size, composition, and thermomechanical treatment.

Cast Iron. Iron ore is reduced to the metallic form in a blast furnace, yielding a product of molten iron saturated in carbon (about 4%). Most commonly, this "hot metal" is immediately processed to steel by a refining process without allowing it to solidify. Occasionally, it is cast into bars; this product is called *pig iron*. Cast iron is made by remelting pig iron and/or scrap steel in a cupola or electric furnace and casting it into molds to the desired shape of the finished part. Cast iron has a much higher carbon content than steel, usually between 2.5% and 3.75%.

Gray Cast Iron. In gray cast iron, the excess carbon beyond that soluble in iron is present as small flake-shaped particles of graphite. The flakes of graphite account for some of the unique properties of gray iron, in particular, its low tensile strength and ductility, its ability to absorb vibrational energy (damping capacity), and its excellent machinability. Cast iron is easy to cast because it has a lower melting point than steel, and the formation of the low-density graphite offsets solidification shrinkage so that minimal dimensional changes occur on freezing. Other elements in the composition of ordinary gray cast iron are important chiefly insofar as they affect the tendency of carbon to form as graphite rather than in chemical combination with the iron as iron carbide (Fe_3C). Silicon is most effective in promoting the formation of graphite. Slower cooling rates during freezing also favor the formation of graphite as well as increase the size of the flakes. Cooling rate also affects the mode of decomposition of the carbon retained in solution during freezing. Slow cooling favors complete precipitation as graphite, leaving a soft ferrite matrix, while fast cooling produces a stronger matrix containing Fe_3C (as pearlite). The tensile strength of gray cast iron typically ranges from 140 to 410 MPa (20,000 to 60,000 lb/in^2). Corresponding compressive strengths are 575 to 1300 MPa (85,000 to 190,000 lb/in^2). Young's modulus may range from 70 to 150 GPa (10×10^6 to 20×10^6 lb/in^2), depending on the microstructure.

White Cast Iron. Careful adjustment of composition and cooling rate can cause all the carbon in a cast iron to appear in the combined form as pearlite or free carbide. This structure is very hard and brittle and has few engineering applications beyond resistance to abrasion. This product does serve, however, as an intermediate product in the production of malleable cast iron described in the following paragraph.

Malleable Cast Iron. By annealing white cast iron at about 950°C, the combined carbon will decompose to graphite. This graphite grows in a spheroidal shape rather than the flake-like shape that forms during the freezing of gray cast iron. Because of this difference in graphite shape, malleable iron is much tougher and stronger. If the castings are slowly cooled from the malleabilizing temperature, the matrix can be converted to ferrite, with all the carbon appearing as graphite; this is a very tough product. Faster cooling will yield a pearlitic matrix with greater strength and hardness. It is also possible to quench and temper malleable iron for optimum combinations of strength and toughness. By careful control of composition, the malleabilizing cycle can be carried out in 8 to 20 h.

Nodular Cast Iron. *Nodular cast iron* has the same carbon content as gray iron; however, the addition of a few hundreds of 1% of either magnesium or cerium causes the uncombined carbon to form spheroidal particles during solidification instead of graphite flakes. Strength properties comparable with those of steel may be achieved in the pearlitic iron. The softer ferritic and pearlitic as-cast irons exhibit considerable ductility, 10% elongation or more. As the hardness and strength are increased by appropriate heat treatment or the thickness of the casting decreased below approximately $^1/_4$ in, the ductility decreases. An austenitic form of nodular iron may be obtained by adding various amounts of silicon, nickel, manganese, and chromium. For many purposes, nodular iron exhibits properties superior to those of either gray or malleable cast iron.

Chilled Cast Iron. *Chilled cast iron* is made by pouring cast iron into a metallic mold which cools it rapidly near the surfaces of the casting, thus forming a wear-resisting skin of harder material than the body of the metal. The rapid cooling decreases the proportion of graphite and increases the combined carbon, resulting in the formation of white cast iron.

Alloy Cast Iron. *Alloy cast iron* contains specially added elements in sufficient amount to produce measurable modification of the physical properties. Silicon, manganese, sulfur, and phosphorus, in quantities normally obtained from raw materials, are not considered alloy additions. Up to about 4% silicon increases the strength of pure iron; greater content produces a matrix of dissolved silicon that is weak, hard, and brittle. Cast irons with 7% to 8% silicon are used for heat-resisting purposes and with 13% to 17% silicon form acid- and corrosion-resistant alloys, which, however, are extremely brittle. Manganese up to 1% has little effect on mechanical properties but tends to inhibit the harmful effects of sulfur. Nickel, chromium, molybdenum, vanadium, copper, and titanium are commonly used alloying elements. The methods of processing or of making the alloy additions to the iron influence the final properties of the metal; hence, a specified chemical analysis is not sufficient to obtain required qualities. Heat treatment is also employed on alloy irons to enhance the physical properties.

Density of Cast Iron. *Density of cast iron* varies considerably depending on the carbon content and the proportion of the carbon that is present as graphite. Using the density of pure iron, 7.86, as a reference, the density of cast iron may range from 7.60 for white cast iron to as low as 6.80 for gray cast iron.

Thermal Properties of Cast Iron. Thermal properties vary somewhat with the composition and the proportions of graphitic carbon. The average specific heat from 20 to 110°C is 0.119; thermal conductivity, 0.40 W/(cm^3)(°C); coefficient of linear expansion, 0.0000106/°C at 40°C.

Values of modulus of elasticity for ferrous metals may be assumed approximately as shown in Table 4-13. The values for all steels are fairly constant, whereas for cast irons the modulus increases somewhat with increased strength of material. Alloy steels have practically the same modulus as plain carbon steels unless large amounts, say, 10%, of alloying material are added; for large percentages of alloying elements, the modulus decreases slightly. The modulus of steels is not affected by heat treatment.

TABLE 4-13 Approximate Modulus of Elasticity for Ferrous Metals

Metal	Modulus in tension-compression, GPa (lb/in^2 × 10^{-6})	Modulus in shear, GPa (lb/in^2 × 10^{-6})
All steels	206 (30.0)	83 (12.0)
Wrought iron	186 (27.0)	75 (10.8)
Malleable cast iron	158 (23.0)	63 (9.2)
Gray cast iron, ASTM No. 20	103 (15.0)	41 (6.0)
Gray cast iron, ASTM No. 60	138 (20.0)	55 (8.0)

Heat Treatment of Steel. The properties of steels can be greatly modified by thermal treatments, which change the internal crystalline structure of the alloy. *Hardening* of steel is based on the fact that iron undergoes a change in crystal structure when heated above its "critical" temperature. Above this critical transformation temperature, the structure is called *austenite*, a phase capable of dissolving carbon up to 2%. Below the critical temperature, the steel transforms to *ferrite*, in which carbon is insoluble and precipitates as an iron carbide compound, Fe_eC (sometimes called *cementite*). If a steel is cooled rapidly from above the critical temperature, the carbon is unable to diffuse to form cementite, and the austenite transforms instead to an extremely hard metastable constituent called *martensite,* in which the carbon is held in supersaturation. The hardness of the martensite depends sensitively on the carbon content. Low-carbon steels (below about 0.20%) are seldom quenched, while steels above about 0.80% carbon are brittle and liable to crack on quenching. Plain carbon steels must be quenched at very fast rates in order to be hardened. Alloying elements can be added to decrease the necessary cooling rates to cause hardening; some alloy steels will harden when cooled in air from above the critical temperature. It should be noted, however, that it is the amount of carbon that primarily determines the properties of the alloy; the alloying elements serve to make the response to heat treatment possible.

Normalizing is a treatment in which the steel is heated over the critical temperature and allowed to cool in still air. The purpose of normalizing is to homogenize the steel. The carbon in the steel will appear as a fine lamellar product of cementite and ferrite called pearlite.

Annealing is similar to normalizing, except the steel is very slowly cooled from above the critical. The carbides are now coarsely divided and the steel is in its softest state, as may be desired for cold-forming or machining operations.

Process annealing is a treatment carried out below the critical temperature designed to recrystallize the ferrite following a cold-working operation. Metals become hardened and embrittled by plastic deformation, but the original state can be restored if the alloy is heated high enough to cause new strain-free grains to nucleate and replace the prior strained structure. This treatment is commonly applied as a final processing for low-carbon steels where ductility and toughness are important, or as an intermediate treatment for such products as wire that are formed by cold working.

Stress-relief annealing is a thermal treatment carried out at a still lower temperature. No structural changes take place, but its purpose is to reduce residual stresses that may have been introduced by previous nonuniform deformation or heating.

Tempering is a treatment that always follows a hardening (quenching) treatment. After hardening, steels are extremely hard, but relatively weak owing to their brittleness. When reheated to temperatures below the critical, the martensitic structure is gradually converted to a ferrite-carbide aggregate that optimizes strength and toughness. When steels are tempered at about 260°C, a particularly brittle configuration of precipitated carbides forms; steels should be tempered above or below this range. Another phenomenon causing embrittlement occurs in steels particularly containing chromium and manganese that are given a tempering cycle that includes holding at or cooling through temperatures around 567 to 621°C. Small molybdenum additions retard this effect, called *temper brittleness*. It is believed to be caused by a segregation of trace impurity elements to the grain boundaries.

Manganese Steels. Manganese is present in all steels as a scavenger for sulfur, an unavoidable impurity; otherwise, the sulfur would form a low-melting constituent containing FeS, and it would be impossible to hot work the steel. The manganese content should be about 5 times the sulfur to provide protection against this "hot shortness." Beyond this amount, manganese increases the hardness of the steel and also has a strong effect on improving response to hardening treatment, but increases susceptibility to temper brittleness. Manganese can be specified up to 1.65% without the steel being classified as an alloy steel. The alloy containing 12% to 14% Mn and around 1% carbon is called *Hadfield's manganese steel*. This alloy can be quenched to retain the austenite phase and is quite tough in this condition. When deformed, this austenite transforms to martensite, which confers exceptional wear and abrasion resistance. Applications for this unique steel include railroad switches, crushing and grinding equipment, dipper bucket teeth, etc.

Vanadium Steels. In amounts up to 0.01%, vanadium has a powerful strengthening effect in microalloyed high-strength, low-alloy steels. In alloy steels, 0.1% to 0.2% vanadium is used as a deoxidizer and carbide-forming addition to promote fine-grained tough steels with deep hardening characteristics. Vanadium accentuates the benefits derived from other alloying elements such as manganese, chromium, or nickel, and it is used in a variety of quaternary alloys containing these elements. Vanadium in amounts of 0.15% to 2.50% is an important element in a large number of tool steels.

Silicon Steels. Silicon is present in most constructional steels in amounts up to 0.35% as a deoxidizer to enhance production of sound ingot structures. Silicon increases the hardenability of steel slightly and also acts as a solid solution hardener with little loss of ductility in amounts up to 2.5%. Silicon in amounts of about 4.5% is a major ingredient in electrical steel sheets. Silicon improves the magnetic properties of iron, but even more important, these steels can be fabricated to produce controlled grain size and orientation. Since permeability depends on crystal orientation, exceptionally small core losses are obtained by using grain-oriented silicon steel in motors and transformers. Alloys containing 12% to 14% Si are exceptionally resistant to corrosion by acids. This alloy is too brittle to be rolled or forged, but it can be cast and is widely used as drainpipe in laboratories and for containers of mineral acids.

Nickel Steels. Nickel is used as a ferrite strengthener and improves the toughness of steel, especially at low temperatures. Nickel also improves the hardenability and is particularly effective when used in combination with chromium. Nickel acts similarly to copper in improving corrosion resistance to atmospheric exposure. Certain iron-nickel alloys have particularly interesting properties and are used for special applications: *Invar* (36% Ni) has a very low temperature coefficient of expansion; *Platinite* (46% Ni) has the same expansion coefficient as platinum; and the 39% Ni alloy has the same coefficient as low-expansion glasses. These alloys are useful as gages, seals, etc. *Permalloys* (45% and 76% Ni) have exceptionally high permeability and are used in transformers, coils, relays, etc.

Chromium Steels. In constructional steels, chromium is used primarily as a hardener. It improves response to heat treatment and also forms a series of complex carbide compounds that improve wear and high-temperature properties. For these purposes, the amount of chromium used is less than 2%. Alloys containing around 5% Cr retain high hardness at elevated temperatures, and have applications as die steels and high-temperature processing equipment. Alloys containing more than 11% Cr have exceptional resistance to atmospheric corrosion and form the basis of the stainless steels.

Stainless Steels. Iron-base alloys containing between 11% and 30% chromium form a tenacious and highly protective chrome oxide layer that gives these alloys excellent corrosion-resistant properties. There are a great number of alloys that are generally referred to as *stainless steels*, and they fall into three general classifications.

Austenitic stainless steels contain usually 8% to 12% nickel, which stabilizes the austenitic phase. These are the most popular of the stainless steels. With 18% to 20% chromium, they have the best corrosion resistance and are very tough and can undergo severe forming operations. These alloys are susceptible to embrittlement when heated in the range of 593 to 816°C. At these temperatures, carbides precipitate at the austenite grain boundaries, causing a local depletion of the chromium content in the adjacent region, so this region loses its corrosion resistance. Use of "extra low carbon" grades and grades containing stabilizing additions of strong carbide-forming elements such as niobium minimizes this problem. These alloys are also susceptible to stress corrosion in the presence of chloride environments.

Ferritic stainless steels are basically straight Fe-Cr alloys. Chromium in excess of 14% stabilizes the low-temperature ferrite phase all the way to the melting point. Since these alloys do not undergo a phase change, they cannot be hardened by heat treatment. They are the least expensive of the stainless alloys.

Martensitic stainless steels contain around 12% Cr. They are austenitic at elevated temperatures but ferritic at low; hence they can be hardened by heat treatment. To obtain a significant response to heat treatment, they have higher carbon contents than the other stainless alloys. Martensitic alloys are used for tools, machine parts, cutting instruments, and other applications requiring high strength. The austenitic alloys are nonmagnetic, but the ferritic and martensitic grades are ferromagnetic.

Heat-Resistant Alloys. Heat-resistant alloys are capable of continuous or intermittent service at temperatures in excess of 649°C. There are a great number of these alloys; they are best considered by class. *Iron-chromium alloys* contain between 10% and 30% chromium. The higher the chromium, the higher is the service temperature at which they can operate. They are relatively low-strength alloys and are used primarily for oxidation resistance. *Iron-chromium-nickel alloys* have chromium in excess of 18%, nickel in excess of 7%, and always more chromium than nickel. They are austenitic alloys and have better strength and ductility than the straight Fe-Cr alloys. They can be used in both oxidizing and reducing environments and in sulfur-bearing atmospheres. Iron-nickel-chromium alloys have more than 10% Cr and more than 25% Ni. These are also austenitic alloys and are capable of withstanding fluctuating temperatures in both oxidizing and reducing atmospheres. They are used extensively for furnace fixtures and components and parts subjected to nonuniform heating. They are also satisfactory for electric resistance-heating elements. *Nickel-base alloys* contain about 50% Ni, and also contain some molybdenum. They are more expensive than iron-base alloys, but have better high-temperature mechanical properties. *Cobalt-base alloys* contain about 50% cobalt and have especially good creep and stress-rupture properties. They are widely used for gas-turbine blades. Most of these alloys are available in both cast and wrought form; the castings usually have higher carbon contents and often small additions of silicon and/or manganese to improve casting properties.

4.4.3 Steel Strand and Rope

Iron and Steel Wire. Annealed wire of iron or very mild steel has a tensile strength in the range of 310 to 415 MPa (45,000 to 60,000 lb/in^2); with increased carbon content, varying amounts of cold drawing, and various heat treatments, the tensile strength ranges all the way from the latter figures up to about 3450 MPa (500,000 lb/in^2), but a figure of about 1725 MPa (250,000 lb/in^2) represents the ordinary limit for wire for important structural purposes. For example, see the following paragraph on bridge wire. Wires of high carbon content can be tempered for special applications such as spring wire. The yield strength of cold-drawn steel wire is 65% to 80% of its ultimate strength. For examples showing the effects of drawing and carbon content on wire, see *Making, Shaping, and Treating of Steel*, U.S. Steel.

Galvanized-Steel Bridge Wire. The manufacture of high-strength bridge wire like that used for the cables and hangers of suspension bridges such as the San Francisco–Oakland Bay Bridge, the Mackinac Bridge in Michigan, and the Narrows Bridge in New York is an excellent example of careful control of processing to produce a quality material. The wire is a high-carbon product containing 0.75% to 0.85% carbon with maximum limits placed on potentially harmful impurities. Rolling temperatures are carefully specified, and the wire is subjected to a special heat treatment called *patenting*. The steel is transformed in a controlled-temperature molten lead bath to ensure an *optimal* microstructure. This is followed by cold drawing to a minimum tensile strength of 1550 MPa

(225,000 lb/in^2) and a 4% elongation. The wire is given a heavy zinc coating to protect against corrosion. Joints or splices are made with cold-pressed sleeves which develop practically the full strength of the wire. Fatigue tests of galvanized bridge wire in reversed bending indicate that the endurance limit of the coated wire is only about 345 to 415 MPa (50,000 to 60,000 lb/in^2).

Wire Rope. Wire rope is made of wires twisted together in certain typical constructions and may be either flat or round. Flat ropes consist of a number of strands of alternately right and left lay, sewed together with soft iron to form a band or belt; they are sometimes of advantage in mine hoists. Round ropes are composed of a number of wire strands twisted around a hemp core or around a wire strand or wire rope. The standard wire rope is made of six strands twisted around a hemp core, but for special purposes, four, five, seven, eight, nine, or any reasonable number of strands may be used. The hemp is usually saturated with a lubricant, which should be free from acids or corrosive substances; this provides little additional strength but acts as a cushion to preserve the shape of the rope and helps to lubricate the wires. The number of wires commonly used in the strands are 4, 7, 12, 19, 24, and 37, depending on the service for which the ropes are intended. When extra flexibility is required, the strands of a rope sometimes consist of ropes, which in turn are made of strands around a hemp core. Ordinarily, the wires are twisted into strands in the opposite direction to the twist of the strands in the rope. The makeup of standard hoisting rope is 6×19; extrapliable hoisting rope is 8×19 or 6×37; transmission or haulage rope is 6×7; hawsers and mooring lines are 6×12 or 6×19 or 6×24 or 6×37, etc.; tiller or hand rope is 6×7; highway guard-rail strand is 3×7; galvanized mast-arm rope is 9×4 with a cotton center. The tensile strength of the wire ranges, in different grades, from 415 to 2415 MPa (60,000 to 350,000 lb/in^2), depending on the material, diameter, and treatment. The maximum tensile efficiency of wire rope is 90%; the average is about 82.5%, being higher for 6×7 rope and lower for 6×37 construction. The apparent modulus of elasticity for steel cables in service may be assumed to be 62 to 83×10^6 kPa (9 to 12×10^6 lb/in^2) of cable section. Grades of wire rope are (from historic origins) referred to as traction, mild plow, plow, improved plow, and extra improved plow steel. The most common finish for steel wire is "bright" or uncoated, but various coatings, particularly zinc (galvanized), are used.

4.4.4 Corrosion of Iron and Steel

Principles of Corrosion. Corrosion may take place by direct chemical attack or by electrochemical (galvanic) attack; the latter is by far the most common mechanism. When two dissimilar metals that are in electrical contact are connected by an electrolyte, an electromotive potential is developed, and a current flows. The magnitude of the current depends on the conductivity of the electrolyte, the presence of high-resistance "passivating" films on the electrode surfaces, the relative areas of electrodes, and the strength of the potential difference. The metal that serves as the anode undergoes oxidation and goes into solution (corrodes).

When different metals are ranked according to their tendency to go into solution, the galvanic series, or electromotive series, is obtained. Metals at the bottom will corrode when in contact with those at the top; the greater the separation, the greater the attack is likely to be. Table 4-14 is such a ranking, based on tests by the International Nickel Company, in which the electrolyte was seawater. The nature of the electrolyte may affect the order to some extent. It also should be recognized that very subtle differences in the nature of the metal may result in the formation of anode-cathode galvanic cells: slight differences in composition of the electrolyte at different locations on the metal surface, minor segregation of impurities in the metal, variations in the degree of cold deformation undergone by the metal, etc. It is possible for anode-cathode couples to exist very close to each other on a metal surface. The electrolyte is a solution of ions; a film of condensed moisture will serve.

Corrosion Prevention. An understanding of the mechanism of corrosion suggests possible ways of minimizing corrosion effects. Some of these include (1) avoidance of metal combinations that are not compatible, (2) electrical insulation between dissimilar metals that have to be used together, (3) use of a sacrificial anode placed in contact with a structure to be protected (this is an expensive technique but can be justified in order to protect such structures as buried pipelines and ship hulls),

TABLE 4-14 Galvanic Series of Alloys in Seawater

↑	Platinum
	Gold
Noble or cathodic	Graphite
	Titanium
	Silver
	[Chlorimet 3 (62 Ni, 18 Cr, 18 Mo)
	Hastelloy C (62 Ni, 17 Cr, 15 Mo)]
	[18-8 Mo stainless steel (passive)
	18-8 stainless steel (passive)
	Chromium stainless steel 11–30% Cr (passive)]
	[Inconel (passive) (80 Ni, 13 Cr, 7 Fe)
	Nickel (passive)]
	Silver solder
	[Monel (70 Ni, 30 Cu)
	Cupronickels (60–90 Cu, 40–10 Ni)
	Bronzes (Cu–Sn)
	Copper
	Brasses (Cu–Zn)]
	[Chlorimet 2 (66 Ni, 32 Mo, 1 Fe)
	Hastelloy B (60 Ni, 30 Mo. 6 Fe, 1 Mn)]
	[Inconel (active)
	Nickel (active)]
	Tin
	Lead
	Lead–tin solders
	[18-8 Mo stainless steel (active)
	18-8 stainless steel (active)]
	Ni-Resist (high Ni cast iron)
	Chromium stainless steel, 13% Cr (active)
	[Cast iron
	Steel or iron]
	2024 aluminum (4.5 Cu, 1.5 Mg, 0.6 Mn)
Active or anodic	Cadmium
	Commercially pure aluminum (1100)
↓	Zinc
	Magnesium and magnesium alloys

Note: Alloys will corrode in contact with those higher in the series. Brackets enclose alloys so similar that they can be used together safely.

Source: Fontana and Green, *Corrosion Engineering*, McGraw-Hill, New York.

(4) use of an impressed *emf* from an external power source to buck out the corrosion current (called *cathodic protection*), (5) avoiding the presence of an electrolyte—especially those with high conductivities, and (6) application of a protective coating to either the anode or the cathode. The problems of corrosion control are complex beyond these simple concepts, but since the use of protective coatings on iron and steel is extensive, this subject is treated in the following sections.

Protective Coatings. Protective coatings may be selected to be inert to the corrosive environment and insolate the base metal from exposure, or the coating may be selected to have reasonable resistance to attack but act sacrificially to protect the base metal. Protective coatings may be considered in four broad classes: paints, metal coatings, chemical coatings, and greases. Painting is commonly used for the protection of structural iron and steel but must be maintained by periodic renewal. Metal coatings take various ranks in protective effectiveness, depending on the metal used and its characteristics as a coating material. A wide variety of metals are used to coat steels: zinc, tin, copper, nickel, chromium, cobalt, lead, cadmium, and aluminum; coatings of gold and silver are also used for decorative purposes. Coatings may be applied by these principal methods: hot dipping, cementation, spraying, electroplating, and vapor deposition. The latter may involve simply evaporation and condensation of the deposited metal or may include a chemical reaction between the vapor and the metal to be coated.

Zinc coatings are more widely used for the protection of structural iron and steel than coatings of any other type. The hot-dip process is the earliest type known and is very extensively used at the present time; two improvements, the Crapo process and the Herman, or "galvannealed," process, are used in galvanizing wire. The cementation, or sherardizing, process consists of heating the articles for several hours in a packing of zinc dust in a slowly rotating container. Electroplating is also employed, and heavier coatings can be obtained than are usual with the hot-dip process, but adherence is difficult to obtain, and this process is not often used. See ASTM specifications for zinc-coated iron and steel products.

Aluminum coatings are applied by a cementation process which is commercially known as *calorizing*. The articles to be coated are packed in a drum in a mixture of powdered aluminum, aluminum oxide, and a small amount of ammonium chloride. The articles are then slowly rotated and heated in an inert atmosphere, usually of hydrogen. Such coatings are very resistant to oxidation and sulfur attack at high temperatures. Aluminum coatings also can be applied by the hot-dipping method and then are heat-treated to improve the alloy bond. Aluminum also can be applied by spraying. Aluminum-coated steel is used extensively for oxidation protection, for example, for heat ducts and automobile mufflers. Aluminum-zinc coatings, applied by hot dipping, have been developed that combine the high-temperature protection of aluminum with the sacrificial protection of zinc.

Almost all *tin coatings* are now applied by electrolytic deposition methods. The accurate control obtained by electrolytic deposition is important because of the high cost of tin. Unlike zinc, tin is electropositive to iron. The coating must remain intact; once penetrated, corrosion of the iron will be accelerated. If a zinc coating is penetrated, the zinc will still sacrifically protect the adjacent exposed iron. Tin has good corrosion resistance, is nontoxic, readily bonds to steel, is easily soldered, and is extensively utilized by the container industry for food and other substances.

The objective of *lead coating* of steel is to obtain an inexpensive corrosion-resistant coating. Lead alone will not alloy with iron; so it is necessary to add *tin* to the lead to obtain a smooth, continuous, adherent coating. Originally, about 25% tin (called *terne metal*) was used, but the tin content has been reduced as the price of tin has increased. Since corrosion protection is less effective in this case terne-coated steel is not used extensively. Applications include uses where corrosion is not too critical or likely, such as gasoline tanks and roofing sheets, or where the lubricating properties of the soft lead surface help forming operations.

Metal-spray coatings are applied by passing metal wire through a specially constructed spray gun which melts and atomizes the metal to be used as coating. The surface to be sprayed must be roughened to afford good adhesion of the deposited metal. Nearly all the commonly used protective metals can be applied by spraying, and the process is especially useful for coating large members or repairing coatings on articles already in place. Sprayed coatings can also be applied that will resist wear and can be used to build up worn parts such as armature shafts and bearing surfaces or to apply copper coatings to carbon brushes and resistors.

Chromium coatings can be applied by cementation or electroplating. In electroplating, the best results are secured by first plating on a base coating of nickel or nickel copper to receive the chromium. The great hardness of chromium gives it important applications for protection against wear or abrasion; it will also take and retain a high polish. Very thin coatings have a tendency to be inefficient as a result of the presence of minute pinholes.

Electroplating is employed in the application of coatings of nickel, brass, copper, chromium, cadmium, cobalt, lead, and zinc. Only cadmium, chromium, and zinc are electronegative to iron. The other metals mentioned are employed because of their own corrosion-resistant properties and because they afford surface finishes having certain desirable characteristics.

Protective Paints. Protective paints are extensively employed to protect heavily exposed structures of iron and steel, such as bridges, tanks, and towers. The protection is not permanent but gradually wears away under weather exposure and must be reviewed periodically. Various specially prepared paints are used for protecting the surface from dampness, oxidizing gases, and smoke. No one paint is suitable for all purposes but the choice depends on the nature of the corrosive influence present. Asphaltum and tar protect the surface by formation of an impervious film. A chemical protective action is exerted by paints containing linseed oil as the vehicle and red lead as the pigment; linseed oil absorbs oxygen from the atmosphere and forms a thick elastic covering, a formation hastened by adding salts of manganese or lead to the oil. All dryers, vehicles, and pigments used in paint must be inert to the steel; otherwise, corrosion will be hastened instead of prevented. Graphite- and aluminum-flake pigments give very impermeable films but do not show the inhibitive action of red lead or zinc when the films are scratched. Aluminum has the advantage of reflecting both infrared and ultraviolet rays of the sun; hence, it protects the vehicle from a source of deterioration and is used to paint gasoline-storage tanks to prevent excessive heating due to the sun's rays. A large number of new protective coatings have been developed recently from synthetic materials such as silicones, artificial rubbers, and phenolic plastics. Many of these are tightly adhering compounds in the form of paints or varnishes which offer rather good protection against a wide variety of chemical attacks. The majority of these new coatings, however, are sensitive to abrasion, and many of them must be baked on to secure full effectiveness.

Corrosion-Resistant Ferrous Alloys. Corrosion-resistant ferrous alloys such as rustless or stainless iron and steel have come into use for both structural and ornamental purposes but on account of their chromium and nickel contents are relatively expensive in comparison with the ordinary structural steels. Copper-bearing iron and steel, containing about 0.15% to 0.25% copper, are used extensively; the copper content tends to retard corrosion slightly but does not prevent it, and some protective coating is usually necessary. Some structural uses have been made of these steels without applying special protective coatings. A tightly adherent brown oxide surface film forms from weathering to serve as the future "protective coating."

Copperweld. A series of steel products, including wire, wire rope, bars, clamps, ground rods, and nails, that contain a copper-clad surface are made by the Copperweld process. The copper coating is intimately bonded to the steel by pouring a ring of molten copper about a heated steel billet fastened in the center of a refractory mold. The solidified composite ingot is then hot-rolled to bar stock and subsequently cold drawn to the various wire sizes. The thickness of the copper coating on wire is 10% to 12½% of the wire radius and produces a high-strength steel wire with a resistance to corrosion similar to that of a solid copper wire. Their increased electrical conductivity over that of a solid steel wire or rod makes the Copperweld products suitable for high-strength conductors, ground rods, aerial cable messengers, etc.

4.4.5 Nonferrous Metals and Alloys

Copper. Numerous commercial "coppers" are available. The standard product is *tough-pitch* copper, which contains about 0.04% oxygen. If it has been electrolytically refined, it is called *electrolytic tough pitch.* This copper cannot be heated in reducing atmospheres because the oxygen will react with hydrogen and severely embrittle the alloy. Various deoxidized varieties are made. When deoxidized with phosphorus, there is some loss of electrical conductivity depending on the amount of residual phosphorus and the extent to which other impurities are reduced and redissolve in the metal. As a general principle, alloying elements that dissolve in copper reduce conductivity sharply; those that are insoluble have little effect.

Copper castings are improved by using special deoxidizers such as Boroflux and silicocalcium copper alloy. By the use of these deoxidizers, the castings are improved structurally, and the electrical conductivity can be increased to about 80% to 90% of standard annealed copper. Boroflux is a mixture of boron suboxide, boric anhydride, magnesia, and magnesium; for data on its use, see publications of the General Electric Company.

Oxygen-free high-conductivity copper is deoxidized with carbon, and thus is free of residual oxide or deoxidizer. It is a more expensive product but does not suffer the potential embrittlement of tough pitch and is capable of more severe cold-forming operations at the cost of a slight loss of electrical conductivity. Free-machining copper contains lead or tellurium that drops conductivity from 3% to 5%. Since copper is a very difficult material to machine, this may be a small sacrifice for certain applications. Small amounts of silver improve resistance to elevated-temperature softening with no loss of physical or mechanical properties.

Brass. Brasses are alloys of copper and zinc; commercial brasses contain from 5% to 45% zinc. A wide variety of properties are obtainable in the brasses. In general, the alloys have excellent corrosion resistance, good mechanical properties, colors ranging from red to gold to yellow to white, and are available in a wide variety of cast and wrought shapes. The alloy of 30% zinc has an optimum combination of strength and ductility. It is called "cartridge brass," since an early application was drawing of cartridge shells. It is the most commonly used brass alloy.

Muntz metal contains nominally 40% zinc and is a two-phase alloy that is readily hot-worked in the high-temperature form and develops good strength when cooled. It is used for extruded shapes and for bolts, fasteners, and other high-strength applications. The properties of brass can be modified by small additions of numerous alloying elements; those commonly used include silicon, aluminum, manganese, iron, lead, tin, and nickel. The addition of 1% tin to cartridge brass results in an alloy called *Admiralty brass*, which has very good corrosion resistance and is extensively used in heat exchangers.

Brasses, especially the high zinc-bearing alloys, are subject to a corrosion phenomenon called *dezincification.* It involves a selective loss of zinc from the surface and the formation of a spongy copper layer accompanied by deterioration of mechanical properties. It is more likely to occur with the high-zinc brasses in contact with water containing dissolved CO_2 at elevated temperatures. Like many other metals, the brasses are susceptible to stress-corrosion craking—an embrittlement due to the combined action of stress and a selective corrosive agent. In the case of brass, the particular agent responsible for stress-corrosion cracking is ammonia and its compounds. Brass products that might be exposed to such environments should be stress-relief annealed before being placed in service. For details on compositions and properties of brasses, see the appropriate ASTM specifications and the publications of brass producers.

Bronze, or Copper-Tin, Alloys. Bronze is an alloy consisting principally of copper and tin and sometimes small proportions of zinc, phosphorus, lead, manganese, silicon, aluminum, magnesium, etc. The useful range of composition is from 3% to 25% tin and 95% to 75% copper. Bronze castings have a tensile strength of 195 to 345 MPa (28,000 to 50,000 lb/in^2), with a maximum at about 18% of tin content. The crushing strength ranges from about 290 MPa (42,000 lb/in^2) for pure copper to 1035 MPa (150,000 lb/in^2) with 25% tin content. *Cast bronzes* containing about 4% to 5% tin are the most ductile, elongating about 14% in 5 in. Gunmetal contains about 10% tin and is one of the strongest bronzes. Bell metal contains about 20% tin. *Copper-tin-zinc alloy* castings containing 75% to 85% copper, 17% to 5% zinc, and 8% to 10% tin have a tensile strength of 240 to 275 MPa (35,000 to 40,000 lb/in^2), with 20% to 30% elongation. *Government bronze* contains 88% copper, 10% tin, and 2% zinc; it has a tensile strength of 205 to 240 MPa (30,000 to 35,000 lb/in^2), yield strength of about 50% of the ultimate, and about 14% to 16% elongation in 2 in; the ductility is much increased by annealing for ½ h at 700 to 800°C, but the tensile strength is not materially affected. *Phosphor bronze* is made with phosphorus as a deoxidizer; for malleable products such as wire, the tin should not exceed 4% or 5%, and the phosphorus should not exceed 0.1%. *United States Navy bronze* contains 85% to 90% copper, 6% to 11% tin, and less than 4% zinc, 0.06% iron, 0.2% lead, and 0.5% phosphorus; the minimum tensile strength is 310 MPa (45,000 lb/in^2), and elongation at least 20% in 2 in. *Lead bronzes* are used for bearing metals for heavy duty; an ordinary composition is 80% copper, 10% tin, and 10% lead, with less than 1% phosphorus. *Steam or valve bronze* contains approximately 85% copper, 6.5% tin, 1.5% lead, and 4% zinc; the tensile strength is 235 MPa (34,000 lb/in^2), minimum, and elongation 22% minimum in 2 in (ASTM Specification B61). The bronzes have a great many industrial applications where their combination of tensile properties and corrosion resistance is especially useful.

Beryllium-Copper Alloys. Beryllium-copper alloys containing up to 2.75% beryllium can be produced in the form of sheet, rod, wire, and tube. The alloys can be hardened by a heat treatment consisting of quenching from a dull red heat, followed by reheating to a low temperature to hasten the precipitation of the hardening constituents. Depending somewhat on the heat treatment, the alloy of 2.0% to 2.25% beryllium has a tensile strength of 415 to 650 MPa (60,000 to 193,000 lb/in^2), elongation 2.0% to 10.0% in 2 in, modulus of elasticity 125 × GPa (18 × 10^6 lb/in^2), and endurance limit of about 240 to 300 MPa (35,000 to 44,000 lb/in^2). An outstanding quality of this alloy is its high endurance limit and corrosion resistance; it can be hardened by heat treatment to give great wear resistance and has high electrical conductivity. Typical applications include nonsparking tools for use where serious fire or explosion hazards exist and many electrical accessories such as contact clips and springs or instrument and relay parts. Beryllium is a toxic substance, and care should be taken to avoid ingesting airborne particles during such operations as machining and grinding. For details on properties and uses of beryllium bronzes, see publications of Kawecki Berylco Industries, New York.

Nickel. Nickel is a brilliant metal which approaches silver in color. It is more malleable than soft steel and when rolled and annealed is somewhat stronger and almost as ductile. The tensile strength ranges from 415 MPa (60,000 lb/in^2) for cast nickel to 795 MPa (115,000 lb/in^2) for cold-rolled full-hard strip; yield strength 135 to 725 MPa (20,000 to 105,000 lb/in^2); elongation in 2 in, 2% when full hard to about 50% when annealed; modulus in tension is about 205 GPa (30 × 10^6 lb/in^2). Nickel takes a good polish and does not tarnish or corrode in dry air at ordinary temperatures. It has various industrial uses in sheets, pipes, tubes, rods, containers, and the like, where its corrosion resistance makes it especially suitable.

The greatest tonnage use of nickel is as an alloying element in steels, principally stainless and heat-resisting steels. There are also a variety of copper-nickel alloys whose main applications are based on their excellent corrosion resistance, for example, condenser tubes. Additions of aluminum and titanium to nickel-base alloys result in age-hardening characteristics, and they can be heat-treated to exceptionally high strengths that are retained to high temperatures. The International Nickel Company publishes an extensive list of bulletins describing the characteristics of nickel and nickel alloys.

Monel Metal. Monel metal is a silvery-white alloy containing approximately 66% to 68% nickel, 2% to 4% iron, 2% manganese, and the remainder copper. It can be cast, forged, rolled, drawn, welded, and brazed, and is easily machined. It melts at 1360°C and has a density of 8.80, coefficient of expansion of 14 × 10^{-6} per degree Celsius, thermal conductivity of 0.06 cgs unit, specific heat of 0.127 cal/(g)(°C), and modulus of 175 GPa (25 × 10^6 lb/in^2). The tensile strength ranges from 450 MPa (65,000 lb/in^2) for cast monel metal to 860 MPa (125,000 lb/in^2) in cold-rolled full-hard strip; yield strength 175 to 725 MPa (25,000 to 115,000 lb/in^2). It is highly resistant to corrosion and the action of seawater or mine waters. The industrial uses for it include many applications where its combination of physical properties and corrosion resistance gives it special advantages.

Magnesium Alloys. The outstanding feature of magnesium alloys is their light weight (specific gravity of about 1.8). Alloys containing thorium and rare-earth additions have been developed that retain good strength at temperatures between 260 and 371°C. The correspondingly high strength/weight ratio makes them particularly useful to the aircraft industry. Less exotic alloys, based mainly on alloying with aluminum (up to 10%) and zinc (up to 6%), still have excellent strengths and are heat-treatable. These alloys have many uses where low density is desired: portable tools, ladders, structural members for trucks and buses, housings, etc. Magnesium alloys are available as castings, forgings, extrusions, and rolled-mill products in a variety of shapes. Their thermal coefficient of expansion is about 0.000029/°C, and their melting point is about 620°C. Tensile strengths of castings range from 145 to 235 MPa (21,000 to 34,000 lb/in^2), yield strengths from 62 to 150 MPa (9,000 to 22,000 lb/in^2), and elongation from 1% to 10% in 2 in. Forged or extruded alloys have tensile strengths of 225 to 300 MPa (33,000 to 43,000 lb/in^2), yield strengths 125 to 205 MPa (18,000 to 30,000 lb/in^2), and elongations of 5% to 17% in 2 in. The Brinell hardness ranges from 35 to 78, and the endurance limit from 40 to 115 MPa (6,000 to 17,000 lb/in^2) depending on the alloy and heat treatment. Since magnesium is highly anodic to other common metals, care must be taken in designing with this metal. Protective coatings are used, and care must be taken to avoid forming galvanic couples. Finely divided magnesium will burn but massive sections are safely melted and welded.

Lead. Lead is a heavy, soft, malleable metal with a blue-gray color; it shows a metallic luster when freshly cut, but the surface is rapidly oxidized in moist air. It can be easily rolled into thin sheets and foil or extruded into pipes and cable sheaths but cannot be drawn into fine wire. Although in an ordinary tensile test lead may develop a tensile strength of 17 MPa (2400 lb/in^2), it may creep at ordinary room temperatures at stresses as low as 0.34 MPa (50 lb/in^2). Owing to this tendency to creep, it may fracture under long-continued load at stresses as low as 5.5 MPa (800 lb/in^2), and the ordinary static tensile properties do not have much significance. The resistance of pure lead to corrosion makes it useful in the form of sheets, pipes, and cable coverings, and large quantities of lead are used in the manufacture of various alloys, particularly in alloys for bearings. Common alloys of lead for cable sheatings contain (approximately) 0.04% Cu, 0.75% Sb, or 0.03% Ca. The greatest use of metallic lead is in the manufacture of storage batteries.

Tin. Tin is a silvery-white, lustrous metal, very soft and malleable and of very low tensile strength. It has a density of about 7.3 and melts at 232°C. In ductility it equals soft steel. The tensile strength varies with the speed of testing. As a metal it has few uses except in sheets, but large quantities of it are used in various industrial alloys. Its chief uses are in tin- and terneplate, solder, babbitt and other bearing metals, brass, and bronze. Tin is very resistant to atmospheric corrosion, and water hardly affects it at all; however, it is electronegative to iron and therefore is not an efficient protective coating under atmospheric exposures.

Zinc. Zinc is a bluish-white metal which has a metallic luster on a new fracture. The density of cast zinc ranges from 7.04 to 7.16. At ordinary temperature it is brittle, but in the range of about 100 to 150°C it becomes malleable and can be rolled into sheets and drawn into wire. At 200°C, it becomes so brittle that it can be pulverized. The tensile strength of cast zinc ranges from about 55 to 95 MPa (8000 to 14,000 lb/in^2) in an ordinary testing-machine test and that of drawn zinc from about 150 to 200 MPa (22,000 to 30,000 lb/in^2); it has a poorly defined proportional limit of about 35 MPa (5000 lb/in^2) and exhibits a certain amount of creep at room temperatures; hence, it may fracture in service under constant stresses below its testing-machine strength. It strongly resists atmospheric corrosion but is readily attacked by acids. The principal industrial uses for it are for galvanizing iron and steel, for plates and sheets for roofing and other applications, and for alloying with copper, tin, and other metals; very large quantities are used in the various types of brass. Next to galvanizing, the greatest use of zinc is in the production of die castings. Because of its moderate melting point, good mechanical properties, and especially because it does not attack steel melting pots and dies, it is the most popular die-casting material (although closely rivaled by aluminum). Zinc alloys for die casting contain some aluminum, copper, and magnesium; all ingredients must be very pure or the casting will have poor corrosion resistance and dimensional stability.

Titanium and Titanium Alloys. Titanium alloys are important industrially because of their high strength-weight ratio, particularly at temperatures up to 427°C. The density of the commercial titanium alloys ranges from 4.50 to 4.85 g/cm^3, or approximately 70% greater than aluminum alloy and 40% less than steel. The purest titanium currently produced (99.9% Ti) is a soft, white metal. The mechanical strength increases rapidly, however, with an increase of the impurities present, particularly carbon, nitrogen, and oxygen. The commercially important titanium alloys, in addition to these impurities, contain small percentages (1% to 7%) of (1) chromium and iron, (2) manganese, and (3) combinations of aluminum, chromium, iron, manganese, molybdenum, tin, or vanadium. The thermal conductivity of the titanium alloys is low, about 15 W/m · K at 25°C, and the electrical resistivity is high, ranging from 54 $\mu\Omega \cdot$ cm for the purest titanium to approximately 150 $\mu\Omega \cdot$ cm for some of the alloys. The coefficient of thermal expansion of the titanium alloys varies from 2.8 to 3.6 $\times$ 10^{-6} per degree Celsius, and the melting-point range is from 1371 to 1704°C for the purest titanium. The tensile modulus of elasticity varies between 100 to 120 GPa (15 to 17 $\times$ 10^6 lb/in^2). The mechanical properties, at room temperature, for annealed commercial alloys range approximately as follows: yield strength 760 to 965 MPa (110,000 to 140,000 lb/in^2); ultimate strength 800 to 1100 MPa (116,000 to 160,000 lb/in^2); elongation 5% to 18%; hardness 300 to 370 Brinell. On the basis of the strength-weight ratio many of the titanium alloys exhibit superior short-time tensile properties as compared with many of the stainless and heat-resistant alloys up to approximately 427°C. However, at the same stress and elevated temperature, the creep rate of the titanium alloys is generally higher than that of the heat-resistant alloys. Above about 482°C, the strength properties of titanium alloys decrease rapidly. The corrosion resistance of the titanium alloys in many media is excellent; for most purposes, it is the equivalent or superior to stainless steel.

Aluminum. Aluminum is an important commercial metal possessing some very unique properties. It is very light (density about 2.7) and some of its alloys are very strong, so its strength-weight ratio makes it very attractive for aeronautical uses and other applications in which weight saving is important. Aluminum, especially in the pure form, has very high electrical and thermal conductivities and is used as an electrical conductor in heat exchangers, etc. Aluminum has good corrosion resistance, is nontoxic, and has a pleasing silvery-white color; these properties make it attractive for applications in the food and container industry, architectural, and general structural fields.

Aluminum is very ductile and easily formed by casting and mechanical forming methods. Aluminum owes its good resistance to atmospheric corrosion to the formation of a tough, tenacious, highly insulating, thin oxide film, in spite of the fact that the metal itself is very anodic to other metals. In moist atmospheres, this protective oxide may not form, and some caution must be taken to maintain this film protection. Although aluminum can be joined by all welding processes, this same oxide film can interfere with the formation of good bonds during both fusion and resistance welding, and special fluxing and cleaning must accompany welding operations.

Commercially pure aluminum (99+%) is very weak and ductile: tensile strength of 90 MPa (13,000 lb/in^2), yield strength of 34.5 MPa (5000 lb/in^2), and shearing strength of 62 MPa (9500 lb/in^2). Extrapure grades (electrical conductor grade) are 99.7+% pure, and are even weaker, but have better conductivity.

Heat Treatment of Aluminum Alloys. Alloys of the 1000, 3000, and 5000 series cannot be hardened by heat treatment. They can be hardened by cold working and are available in annealed (recrystallized) and cold-worked tempers. The 5000-series alloys are the strongest non-heat-treatable alloys and are frequently used where welding is to be employed, since welding will generally destroy the effects of hardening heat treatment. The remaining wrought alloys can be hardened by controlled precipitation of alloy phases. The precipitation is accomplished by first heating the alloy to dissolve the alloying elements, followed by quenching to retain the alloy in supersaturation. The alloys are then "aged" to develop a controlled size and distribution of precipitate that produces the desired level of hardening. Some alloys naturally age at room temperature; others must be artificially aged at elevated temperatures.

4.4.6 Stone, Brick, Concrete, and Glass Brick

Building Stone. Stone is any natural rock deposit or formation of igneous, sedimentary, and/or metamorphic origin, in either its original or its altered form. Building stone is the quarried product of such deposit or formation, which is suitable for structural and ornamental purposes. Igneous or volcanic rock, such as granite or basalt, is rock of plutonic or volcanic origin, formed from a fused condition and crystalline in structure. Sedimentary rock, such as limestone, dolomite, and sandstone, is formed by the deposition of particles from water and laminated in structure. Metamorphic rock, such as gneiss, marble, and slate, is rock formation which, in the natural ledge, has undergone marked change in microstructure or character due to heat, pressure, or moisture and therefore exists in form different from the original.

Portland Cement. *Portland cement* is produced by sintering a proportional mixture of lime and clay, which is subsequently ground with the addition of gypsum (to retard the rate of setting). The properties of the clay and limestone determine the principal characteristics: fineness, soundness, time of set, and strength. Strength is measured from briquettes made of a mortar of 1 part cement to 3 parts sand, and measured under specified conditions. Minimum tensile strengths are 2 MPa (275 1b/in^2) at 7 days, and 2.6 MPa (350 1b/in^2) at 28 days.

Mortar. *Mortar* is a mixture of sand, screenings, or similar inert particles with cement and water, which has the capacity of hardening into a rocklike mass. The inert particles are usually less than $1/4$ in in size. The proportions of cement to sand range all the way from 1:0 to 1:4 for various purposes.

Concrete. *Concrete* is a mixture of crushed stone, gravel, or similar inert material with a mortar. The maximum size of inert particles is variable but usually less than 2 in. These inert constituents of mortar and concrete are known as the *aggregate.* In making good concrete, the properties of the aggregate are as important as those of the cement. The fine aggregates consist of sand, screenings, mine tailings, pulverized slag, etc., with particle sizes less than $1/4$ in; the coarse aggregates consist of crushed stone, gravel, cinders, slag, etc. Rubble concrete is made by embedding a considerable proportion of boulders or stone blocks in concrete. The proportions by volume of cement, sand, and coarse aggregate range all the way from 1:1:2 for high compressive strength to 1:4:8 for structures requiring mass more than strength. In general, the strength of concrete increases with the density and richness of mix (proportion of cement) but is decreased in proportion to the amount of mixing water that is added beyond that required to produce a plastic workable mixture. In controlling the quality of a concrete of given mix, the ratio of the volume of mixing water to the volume of cement (water-cement ratio) is often used as a criterion of the strength. For proper curing, the concrete should be kept moist for at least a week after placing, and care should be taken to prevent its freezing in cold weather during the early stages of curing. Freshly poured concrete gains strength very slowly in cold weather. Various admixtures are often added in small amounts to modify pouring characteristics and setting time as well as physical characteristics such as resistance to freezing and thawing cycles, wear, abrasion, and permeability.

For concrete of common proportions cured under good conditions, 25% to 40% of the 2-year strength is developed in 7 days, 50% to 65% in 1 month, and 70% to 90% in 6 months. The tensile strength of concrete is very low (about one-tenth its compressive strength), and hence in structural members the concrete is usually designed to resist the compressive stresses only, the tensile strength of the concrete being considered negligible. For flexural members, steel reinforcing bars are usually inserted on the tensile side of the beam to resist the tensile stresses.

4.5 WOOD PRODUCTS

By PHILIP MASON OPSAL

Wood Scientist—Consultant, Wood Science, LLC. This author's perception at the outset of the effort to construct this Section is that very few, if any, engineers of any specialty in the last 20 years have received any course work dealing with the subject of wood materials as a construction material. This is due to a number of reasons, including the evolution of course requirements at institution of higher learning to include newer and more modern construction materials, and the advent of the computer causing the need to use significant amounts of time for classes devoted to the teaching of engineers to be computer-wise, effectively providing the basic skills to accomplish all sorts of engineering design functions modeling, simulations, etc., significantly reducing time required for those tasks and increasing productivity for the professional engineer. It is thus that this Chapter is designed to be very basic in wood, not aiming at making professional electrical engineers into wood scientists, but rather to assist in the building of a knowledge base in wood to enable a better understanding of the material made in trees by natural forces. It is therefore purposeful that the bibliography has a considerable array of the old proven and very reliable textbooks of wood science/wood technology to allow the lessor experienced electrical engineer without such experience and/or even without any prior training that even opened up the subject of wood material. For those electrical engineers with more experience and knowledge, the best way to consider proceeding with this chapter to increase and/or enhance the current knowledge base may well be to reference the more contemporary of the references shown and then pursue a course looking for keywords of phrases through the listed references that are to be found in each of those references following along possible avenues to larger more detailed and technical sources to the more current research.

4.5.1 Sources/Trees

The material known as *wood* has as its source *trees*. Natural in origin, the quality of this construction material is based primarily on the availability of the factors that contribute to tree growth, viz, sunlight, moisture, and nutrients. The best quality wood fiber results when these basic factors are at their optimum quantities. Wood materials of the best strength result from the selection of quality species, good forest management during the growing cycle, selective harvesting of only acceptable trees for a use such as utility poles, and then making certain that none of those hand-picked trees are damaged during any of the handling, hauling, and/or processing methods of operations. There are many species of trees present on the earth and on a broad basis are separated into *hardwoods* and *softwoods;* but these names do not truly reflect the relative "hardness" or "softness" of the wood resulting from the trees in either group inasmuch as balsa wood is in the hardwood category and is very lightweight and very weak wood, while a species such as *Douglas-fir*, a wood used in very large quantities as a strong structural member, is classified as a softwood. Generally, the hardwoods have leaves of fairly good size, such as maples, oaks, and elms, and are those trees which have a spreading configuration and are seen typically in yards, parks, and in some cities, lining the streets and avenues where in those locations and for such uses they appear as what are termed *ornamentals*. In hardwood forest stands, however, where, due to the density of the forest, with many trees per acre, natural pruning tends to occur and the tree trunks then grow more pole like with less branch remnants, allowing less deviated grain to from in the outer wood of the tree trunk as growth proceeds in height and circumference. Such was true of the American Chestnut, *Castenea dentata*, from which long tall poles were harvested and subsequently fabricated for use in utilities as supporting structures for conductor, crossarms, insulators, transformers, guys and anchors, regulators, OCRs, etc. Of course chestnut was prized and processed into wood products for many other applications as well. Unfortunately, the chestnut blight found its way from Europe and this organism severely devastated the majority of trees in the United States. Chestnut trees were known for their resistance to decay due to the heartwood section containing a high content of the organic chemicals called *tannins*.

Softwood are also known as *conifers* and most are evergreens keeping their leaves, which are usually "needle-like", all year long (exceptions are Eastern larch and Western larch, respectively, *Larix laricina* and *Larix occidentalis*, which drop their leaves in the fall and grow new leaves each spring). The normal "form class" of the softwood trees is a long tapering trunk extending up to a crown of branches, which originate each year at the top of the tree in a "whorl" or sort of in a circle or sometimes a helical pattern around the circumference of the tree. Knots (earlier in the life of the tree its branches) cause the grain to deviate by the growth of the tree having to grow around the circular or elliptically shaped pattern of braches. It must be remembered that straight grain is the strongest while the greater the deviation of the grain is from straight the weaker the wood is. The configuration of knot whorls in clusters of closely adjacent knots can form a significant weakened section of a wood pole and, as well, weaken structural timbers such as crossarms and crossbraces. It is extremely important to require all wood products to be 100% inspected, examined, and assessed to ensure the receipt at destination of only quality wood capable of safely functioning to support the structure for which their use was designed by the professional electrical engineer.

4.5.2 Wood Structure

Gross wood structure may be used to differentiate between the hardwood and softwood classes and to identify species within each group. The cross section of a log shows several well-defined features from the outer bark, through the wood, to the central pith. The purpose of the bark is to protect the inner living tissues from injury. The inner portion of the bark is a conductive tissue that transports foodstuffs from the leaves to the living cells. The wood portion of the tree has outer sapwood and inner heartwood regions, which often are distinguishable by a difference in color. The lighter-colored sapwood is a living tissue. The darker-colored heartwood consists of entirely dead cells and serves primarily for mechanical support in the tree. The heartwood contains deposits of gums, oils, and other organic infiltrations in the cells. These materials are called *extractives* and import the darker color (ranging from light brown to black) to heartwood. Other differences include durability and permeability. Sapwood of all species is readily destroyed by insects and the decay fungi. The heartwoods of many species are also nondurable, but others are very durable (e.g., cypress, the cedars, and redwood). Sapwood is usually more permeable to liquids, owing largely to the absence of extractives. For this reason, sapwood is more easily treated with preservatives. Wood near the center of the tree is often characterized by fast growth and is termed *juvenile wood*. The properties of juvenile wood are inferior to those of normal heartwood. The strength properties of heartwood and sapwood are the same.

Annual Rings. In temperate zones, the tree increases in diameter and height by division of cells in a singular cell layer, called the *cambium*, capable of dividing and simultaneously supplying a new cell both in the direction of the bark and in the direction of the mass of the tree. The growth increment for each year is called the *annual ring*. The trees of North America have well-defined annual rings due to the production of cellular structure differing during the spurt of growth in the spring, when the tree growth surges and the wood cells produced by such fast growth are very thin-walled and are structurally weaker than those produced later in the growing season as the growth slows and the cells produced tend to be thicker-walled and thus are stronger structurally. These growth patterns have been termed *earlywood* (*Springwood*) and/or *latewood (Summerwood)* and show up as distinctive sections in the annual ring produced in each growing season. Because of this, the age of the tree can be determined by simply counting the rings beginning just inside the bark of the tree and continuing into the pith, the name given to the very center of each tree. (Note: Inasmuch as trees do not always grow in symmetry with purely concentric rings, the growth on two or more sides of the tree may have to be averaged.) Counting the rings and determining the percent of the latewood in the rings is often used as a tool in an attempt to roughly define the strength of wood to be used in a structural capacity. Tropical woods, producing growth nearly continuously, do not usually display seasonal growth rings.

Minute Wood Structure. Minute wood structure differs for hardwood and softwood species. The conifers are composed primarily of long hollow fibers oriented with their long axis parallel to the length of the tree. The ends of these fibers are tapered and the adjacent ends overlapping. The size, shape, and structure of these fibers vary considerably from one species to another, which accounts for much of the variation in properties of the respective wood. Given wood's O*rthotropic* nature, having three distinctly different axes, the longitudinal, the radial, and the tangential, the characteristics of each such as shrinkage are significantly different with a very small amount of shrinkage longitudinally followed by the radial and then by the tangential exhibiting the greatest amount, respectively less than 1% approximately 0.1% to 0.2%, radial 2% to 8%, and tangentially 4% to 14%. The individual fibers, primarily tracheids, are composed of multilayer cell wall enclosed within a central hollow void, much like a soda straw.

Chemical Composition. Chemically, wood consists of approximately 70% cellulose, 25% lignin, and about 5% extractives. The strength of wood may be attributed almost entirely to the cellulose and lignin present in the cell wall. The extractives do not contribute to the cell wall structure but do contribute to such properties as color, odor, taste, and resistance to decay.

Specific Gravity. Specific gravity is a measure of the amount of material contained in a piece of wood. It is calculated by dividing the weight of a given volume of wood by the weight of an equal volume of water. Specific gravity varies according to the amount of water present in wood. For this reason, the ovendry weight of wood is usually used for reported values of specific gravity. The specific gravity of some commercially imported woods is listed in Table 4-15. For clear, straight-grained

TABLE 4-15 Specific Gravity and Shrinkage of Common Woods

Species	Average specific gravity*	Shrinkage, %†	
		Radial	Tangential
Douglas fir, various regions	0.43–0.48	3.6–5.0	6.2–7.8
White ash	0.60	4.8	7.8
Aspen	0.38	3.3	7.9
Yellow birch	0.55	7.2	9.2
White fir	0.37	3.2	7.1
Western red cedar	0.33	2.4	5.0
Northern white cedar	0.31	2.2	4.9
Western hemlock	0.42	4.3	7.9
Southern yellow pines	0.51–0.61	4.4–5.5	7.4–7.8
Tamarack	0.53	3.7	7.4
White oak species	0.63–0.68	4.1–5.5	7.2–10.8
Red oak species	0.59–0.69	4.0–5.5	8.2–10.6
Hickory species	0.69–0.75	7.0–7.8	10.0–12.6

*Weight ovendry and volume at 12% moisture content.
†From green to ovendry condition, based on green dimension.
Source: U.S. Forest Products Laboratory.

wood at a known moisture content, specific gravity is positively correlated with several important properties, including strength and stiffness in bending, tension, and compression. Approximate functions for predicting the mechanical properties of wood for known average and/or specific gravity ranges are given in the *Wood Handbook*. When wood contains natural growth characteristics and/or imperfections, the relationship between properties and specific gravity may be less pronounced due to the disruption of the structural continuity and/or integrity of straight-grained wood known to be the strongest.

4.5.3 Moisture in Wood

Wood is a hygroscopic material. Moisture in wood occurs in three forms: water vapor in airspaces in the cell cavities, capillary water in the cell cavities, and water molecules bound to the hydroxyl groups of the cellulose in the cell wall. In most end-use conditions, when wood is not in contact with water, nearly all the moisture present is bound water and is usually between 3% and 30% of the dry weight of the wood. Since this bound water tends to be at equilibrium with the vapor pressure of the surrounding atmosphere, the maximum amount of bound water in wood occurs in a saturated atmosphere. Any increase in moisture content above this maximum is due to capillary water, acquired from contact with liquid water.

The moisture content of wood is expressed as a percentage of the ovendry weight of wood. It can be measured by weighing a wood sample before and after drying to constant weight at 103°C ± 3°C (217.4°F) using the relationship:

$$\text{Moisture content,\%} = \frac{\text{moist weight–dry weight}}{\text{dry weight}} \times 100$$

When moist wood dries, the liquid water present in the cell capillaries evaporates before the bound water leaves the cell wall. The *fiber-saturation point* is defined as the moisture content of wood at that point at which the cell walls are saturated and there is no free water in the cell cavities. It is at this point that the moisture content is approximately 30%, varying somewhat among species.

Volumetric Changes. Below the fiber-saturation point, water that evaporates from wood results in a reduction in wood volume in the percentages cited earlier in the three directions, longitudinal, radial, and tangential. The amount of volumetric change is positively related to the changes in moisture content and density and due to the orthotropic nature of the wood results in unequal shrinkage. In general, the shrinkage increases with density. Some representative shrinkage values are given Table 4-15. In round wood members, shrinkage results as the wood dries internal stresses develop until reaching the point that relief occurs manifested in the wood checks and splits taking place as to serve to eliminate the stresses. As wood poles dry out, shrinkage in circumference approximates 1% to 2% depending on size.

Wood Seasoning. Most wood products are dried prior to use to remove the large amount of moisture present in freshly cut wood. Wood that has been dried offers a number of advantages including reduced weight and shrinkage, an increase in strength, and durability compared to green wood. Drying may be accomplished by one of several procedures. Two of the most common are air drying and kiln drying. A number of defects may develop during drying if the process is not carefully controlled. These defects are the result of drying stresses due to unequal shrinkage. Most kiln-drying procedures include moisture-equalizing and conditioning treatments to improve moisture uniformity throughout the thickness of the wood product, and to relieve residual stresses. Improper drying may result in warping, checking, or more severe defects.

4.5.4 Thermal Properties of Wood

Temperature affects several properties of wood. As wood is heated, it expands. The coefficient of thermoexpansion for wood averages near 1.1×10^{-6} per degree Celsius for most native species. Wood is a good insulator and does not respond very fast to a change in environmental temperature. The coefficient of thermoconductivity for wood ranges from 0.4 to 0.7 Btu/(h)(°C) for a 1 ft^2 area 1 in in thickness at a moisture content of 12%. The thermoconductivity of wood increases with increasing specific gravity and moisture content.

4.5.5 Electrical Properties of Wood

Three important electrical properties of wood are resistivity, dielectric constant, and power factor. Wood is an excellent insulator. The resistivity of dry wood to the flow of direct current is high, approximately $3 \times 10^{17}\ \Omega \cdot \text{cm/cm}^3$ parallel to the grain. The presence of moisture lowers resistivity. The dielectric constant for wood determines the amount of stored electric potential energy when it is placed in a high-frequency alternating current. The dielectric constant for wood varies over a range from 2.0 for dry wood to 8 for wood above the fiber-saturation point. The dielectric constant is affected by density and grain direction. The power factor determines the amount of energy that is dissipated as heat when wood absorbs power in a high-frequency dielectric field. The power factor for wood is about 2% to 6% at low moisture contents for frequencies between 2 and 15 Hz.

4.5.6 Strength of Wood

Effect of Moisture on Strength. Clear wood generally increases in strength and stiffness as it dries below the fiber-saturation point. The change in strength resulting from a 1% change in moisture content, is approximately 4 for modulus of rupture, 2 for modulus of elasticity, and 5 for compression parallel to the grain.

For structual lumber, the increase in strength of the clear wood is partly offset by the development of seasoning defects such as checks and splits. For this reason, the properties in the green condition are generally used as the base for the development of design stresses for wood. For lumber that is nominally 5 cm (2 in) in thickness, the design stress is increased by up to 25% in bending, 20% in modulus of elasticity, and 37.5% in compression parallel to grain when the moisture content is at or less than 15%.

Effect of Temperature on Strength. In general, heating reduces and cooling increases the mechanical properties of wood. The change is immediate and irreversible for temperatures remaining above 66°C(150°F) for any appreciable period of time. The adverse effect of high temperature is more pronounced at high moisture contents. For elevated temperatures below 66°C(150°F), the immediate loss of strength is recovered when the wood is cooled to ambient conditions. When wood is repeatedly exposed to high temperature, the adverse effect on properties is cumulative leading to creep and ultimate failure without any in ovease in loading.

Mechanical Properties. The mechanical properties of the commercially important woods of the United States have been evaluated in accordance with ASTM Standard D143, which specifies small, clear test specimens to eliminate the influence of naturally occurring physical defects in the wood. Tables 4-16 and 4-17 show some mechanical properties for wood in the green condition and at 12% moisture content, respectively. The green properties are obtained from specimens at essentially the same moisture content as in the living tree, well above the fiber-saturation point. The data in these tables include several of the more important hardwood and softwood species and the mechanical properties of each, which are likely to be uniquely important for specific uses encountered in electrical engineering applications. Adjustments to the values of strength determined for small, clear green test specimens are shown in ASTM Standard D245.

4.5.7 Decay and Preservatives

Decay and Its Prevention. At ordinary temperatures, wood is very stable and, unless attacked by living organisms, remains the same for centuries, either in air or under water. Fungi are the chief enemies of wood, and they thrive best with warmth and abundance of moisture and air, for example, in contact with the ground. Higher temperatures near the surface of the ground, together with adequate air and a greater prevalence of fungi, cause decay to progress faster near the ground line than at several feet below. Proper seasoning, together with protection against the entrance of moisture and impregnating with fungus-inhibiting compounds, which prevent fungi from feeding on the wood, is the best means of preservation. The heartwood of some species is resistant to decay whereas the sapwood of all species is nondurable to decay and insects and the preservative treatment, if done properly, substantially increases the safe service life of wood members.

Wood Preservatives. For a nominal cost, the service life of wood can be greatly increased by the use of preservative treatment. Creosote is extensively used for the protection of poles and crossarms. Southern yellow pine, because of its thick sapwood, requires a pressure treatment. Species with intermediate sapwood thickness such as Douglas-fir, lodgepole pine, and jack pine are treated by either pressure or nonpressure processes. Thin sapwood species such as western red cedar and larch are generally treated by nonpressure processes. The American Wood-Preservers' Association Standards are used to specify preservative chemicals and treatment methods for wood products. The AWPA introduced the "Use-Category System" in 1999, where selection of products based on use and relative exposures to environmental conditions become the main bases of choice of products. After a period of 5 years of experience with the "Use-Category System", the 2004 AWPA Standards have been completely rearranged to offer a more discerning, practical, and understood array of products fitted and custom-tailored by the industry to usage.

Wood preservatives fall into two main classes: (1) oil-borne preservatives and (2) water-borne metallic salts. The former may be further subdivided into (a) coal-tar creosote with and without the mixture of cheaper materials such as petroleum or coal tar and (b) solutions of toxic organic chemicals such as pentachlorophenol dissolved in petroleum oils. Oil-type preservatives are used extensively for products that are exposed to ground contact, whereby resistance to leaching is an important requirement of the preservative. These products include poles, crossties, crossarms, crossbraces, piling, bridge timbers, fence posts, and highway barrier posts. Water-borne preservatives are used mainly for the treatment of lumber. Wood treated with a water-borne preservative is clean, paintable, and odorless.

Creosote is a distillate of coal tar formed during the coking of coal during the steel-making process. Creosote has historically been a most important preservative significant in the preservation of wood members placed in contact with the ground, such as railroad ties, utility poles, piles, and timbers, for application to structures in coastal waters where significant destruction is inflicted on wharves and docks by marine organisms.

TABLE 4-16 Mechanical Properties of Various Woods in the Green Condition Grown in the United States

Species	Moisture content, %	Specific gravity*	Static bending: Modulus of rupture, lb/in²	Static bending: Modulus of elasticity, 1,000 lb/in²	Compression parallel to grain maximum crushing strength, lb/in²	Compression perpendicular to grain stress at proportional limit, lb/in²	Tension perpendicular to grain maximum tensile strength, lb/in²	Hardness†: End, lb	Hardness†: Side, lb	Maximum shearing strength parallel to grain, lb/in²
Ash, black	85	0.45	6,000	1,040	2,300	350	490	590	520	860
Ash, white	42	0.55	9,600	1,460	3,990	670	590	1,010	960	670
Aspen	94	0.35	5,100	860	2,140	180	230	280	300	660
Basswood	105	0.32	5,000	1,040	2,220	170	280	290	250	600
Beech	54	0.56	8,600	1,380	3,550	540	720	970	850	1,290
Birch, yellow	67	0.55	8,300	1,500	3,380	430	430	810	780	1,110
Cottonwood, eastern	111	0.37	5,300	1,010	2,280	200	410	380	340	680
Elm, American	89	0.46	7,200	1,110	2,910	360	590	680	620	1,000
Elm, slippery	85	0.48	8,000	1,230	3,320	420	640	750	660	1,110
Hickory, shagbark	60	0.64	11,000	1,570	4,580	840				1,520
Locust, black	40	0.66	13,800	1,850	6,800	1,160	770	1,640	1,570	1,760
Maple, silver	66	0.44	5,800	940	2,490	370	560	670	590	1,050
Maple, sugar	58	0.56	9,400	1,550	4,020	640		1,070	970	1,460
Oak, red	80	0.56	8,300	1,350	3,440	610	750	1,060	1,000	1,210
Oak, white	68	0.60	8,300	1,250	3,560	670	770	1,120	1,060	1,250
Sweetgum	115	0.46	7,100	1,200	3,040	370	540	670	600	990
Sycamore	83	0.46	6,500	1,060	2,920	360	630	700	610	1,000
Yellow poplar	83	0.40	6,000	1,220	2,660	270	510	480	440	790
Baldcypress	91	0.42	6,600	1,180	3,580	400	300	440	390	810

TABLE 4-16 Mechanical Properties of Various Woods in the Green Condition Grown in the United States (*Continued*)

Species	Moisture content, %	Specific gravity*	Static bending		Compression parallel to grain maximum crushing strength, lb/in²	Compression perpendicular to grain stress at proportional limit, lb/in²	Tension perpendicular to grain maximum tensile strength, lb/in²	Hardness†		Maximum shearing strength parallel to grain, lb/in²
			Modulus of rupture, lb/in²	Modulus of elasticity, 1,000 lb/in²				End, lb	Side, lb	
Cedar, northern white	55	0.29	4,200	640	1,990	230	240	320	230	620
Cedar, Port Orford	43	0.40	6,200	1,420	3,130	280	180	460	400	830
Cedar, western red	37	0.31	5,100	920	2,750	270	230	430	270	710
Douglas-fir, coast‡	38	0.45	7,700	1,560	3,780	380	300	570	500	900
Fir, white	110	0.37	5,900	1,160	2,900	280	300	410	340	760
Hemlock, western	77	0.42	6,600	1,310	3,360	280	290	500	410	860
Larch, western	58	0.48	7,700	1,460	3,760	400	330	580	510	870
Pine, lodgepole	65	0.38	5,500	1,080	2,610	250	220	320	330	680
Pine, ponderosa	91	0.38	5,100	1,000	1,940	280	310	310	320	700
Pine, loblolly	81	0.47	7,300	1,410	3,490	390	260	420	450	850
Pine, longleaf	62	0.54	8,700	1,600	4,300	480	330	550	590	1,040
Pine, shortleaf	81	0.46	7,300	1,390	3,430	350	320	410	440	850
Pine, western white	54	0.36	5,200	1,170	2,650	240	260	310	310	640
Spruce, Engelmann	80	0.32	4,500	960	2,190	220	240	310	260	590
Spruce, Sitka	42	0.37	5,700	1,230	2,670	280	250	430	350	760

Note: 1 lb/in² = 6.895 kPa; 1 lb = 0.4536 kg.
*Specific gravity based on green volume and ovendry weight.
†Load required to embed a 0.444-in ball to half its diameter.
‡Coast Douglas-fir is defined as that coming from counties in Oregon and Washington west of the summit of the Cascade Mountains. For Douglas-fir from other sources, see Western Wood Density Survey, *U.S. Forest Service Res. Paper* FPL 27.

TABLE 4-17 Mechanical Properties of Various Woods in the Air-Dry Condition Grown in the United States

Species	Moisture content, %	Specific gravity*	Static bending: Modulus of rupture, lb/in^2	Static bending: Modulus of elasticity, 1,000 lb/in^2	Compression parallel to grain maximum crushing strength, lb/in^2	Compression perpendicular to grain stress at proportional limit, lb/in^2	Tension perpendicular to grain maximum tensile strength, lb/in^2	Hardness†: End, lb	Hardness†: Side, lb	Maximum shearing strength parallel to grain, lb/in^2
Ash, black	12	0.49	12,600	1,600	5,970	760	700	1,150	850	1,570
Ash, white	12	0.60	15,400	1,770	7,410	1,160	940	1,720	1,320	1,160
Aspen	12	0.38	8,400	1,180	4,250	370	260	510	350	850
Basswood	12	0.37	8,700	1,460	4,730	370	350	520	410	990
Beech	12	0.64	14,900	1,720	7,300	1,010	1,010	1,590	1,300	2,010
Birch,yellow	12	0.62	16,600	2,010	8,170	970	920	1,480	1,260	1,880
Cottonwood, eastern	12	0.40	8,500	1,370	4,910	380	580	580	430	930
Elm, American	12	0.50	11,800	1,340	5,520	690	660	1,110	830	1,510
Elm, slippery	12	0.53	13,000	1,490	6,360	820	530	1,120	860	1,630
Hickory shagbark	12	0.72	20,200	2,160	9,210	1,760				2,430
Locust black	12	0.69	19,400	2,050	10,180	1,830	640	1,580	1,700	2,480
Maple, silver	12	0.47	8,900	1,140	5,220	740	500	1,140	700	1,480
Maple, sugar	12	0.63	15,800	1,830	7,830	1,470		1,840	1,450	2,330
Oak, red	12	0.63	14,300	1,820	6,760	1,010	800	1,580	1,290	1,780
Oak, white	12	0.68	15,200	1,780	7,440	1,070	800	1,520	1,360	2,000
Sweetgum	12	0.52	12,500	1,640	6,320	620	760	1,080	850	1,600
Sycamore	12	0.49	10,000	1,420	5,380	700	720	920	770	1,470
Yellow poplar	12	0.42	10,100	1,580	5,540	500	540	670	540	1,190

TABLE 4-17 Mechanical Properties of Various Woods in the Air-Dry Condition Grown in the United States (*Continued*)

Species	Moisture content, %	Specific gravity*	Static bending: Modulus of rupture, lb/in²	Static bending: Modulus of elasticity, 1,000 lb/in²	Compression parallel to grain maximum crushing strength, lb/in²	Compression perpendicular to grain stress at proportional limit, lb/in²	Tension perpendicular to grain maximum tensile strength, lb/in²	Hardness†: End, lb	Hardness†: Side, lb	Maximum shearing strength parallel to grain, lb/in²
Baldcypress	12	0.46	10,600	1,440	6,360	730	270	660	510	1,000
Cedar, northern white	12	0.31	6,500	800	3,960	310	240	450	320	850
Cedar, Port Orford	12	0.42	11,300	1,730	6,470	620	400	730	560	1,080
Cedar, western red	12	0.33	7,700	1,120	5,020	490	220	660	350	860
Douglas-fir, coast‡	12	0.48	12,400	1,950	7,240	800	340	900	710	1,130
Fir, white	12	0.39	9,800	1,490	5,810	530	300	780	480	1,100
Hemlock, western	12	0.45	11,300	1,640	7,110	550	340	900	540	1,250
Larch, western	12	0.52	13,100	1,870	7,640	930	430	1,120	830	1,360
Pine, lodgepole	12	0.41	9,400	1,340	5,370	610	290	530	480	880
Pine, ponderosa	12	0.40	9,400	1,290	5,320	580	420	570	460	1,130
Pine, loblolly	12	0.51	12,800	1,800	7,080	800	470	750	690	1,370
Pine, longleaf	12	0.58	14,700	1,990	8,440	960	470	920	870	1,500
Pine, shortleaf	12	0.51	12,800	1,760	7,070	810	470	750	690	1,310
Pine, western white	12	0.38	9,500	1,510	5,620	440		440	370	850
Spruce, Engelmann	12	0.34	8,700	1,280	4,770	470	350	560	350	1,030
Spruce, Sitka	12	0.40	10,200	1,570	5,610	580	370	760	510	1,150

Note: 1 lb/in² = 6.895 kPa; 1 lb = 0.4536 kg.
*Specific gravity based on green volume and ovendry weight.
†Load required to embed a 0.444-in ball to half its diameter.
‡Coast Douglas-fir is defined as that coming from counties in Oregon and Washington west of the summit of the Cascade Mountains. For Douglas-fir from other sources, see Western Wood Density Survey, *U.S. Forest Service Res. Paper* FPL 27.

Pentachlorophenol, appearing in crystalline form much like sugar and/or table salt must be dissolved in aromatic-type petroleum oils for best solubility to accomplish movement into the wood cells during the treating process. Low-boiling solvents like mineral spirits are used as a carrier for penta when cleanliness is required for products such as millwork where no contact with the ground is anticipated. Water repellents are added to millwork treatments to both minimize dimensional changes as well as to provide greater stability and permanence to the treated products. Other oil-borne preservatives such as copper napthenate, copper-8-quinolinolate, tri-butylin oxide, and other recent additional are commercially in use.

Water-borne preservatives are generally mixtures of several inorganic salts, the most important of which are salts of copper, chromium arsenic and zinc. Boron is an inorganic water becoming more widely used some of these preservatives give good protection to wood exposed to wet conditions while others are not recommended for such use. Wood treated with water-borne preservatives is clear and paintable after drying to below 25% moisture content.

Several supplemental wood preservatives are used to bolster the original wood preservatives of in-service wood poles, which over time have changed due to leaching and oxidation and thus no longer are protecting those poles as occurred during the earlier years in service. These supplemental preservatives come in the form of (1) bandages containing grease formulations for exterior applications, (2) liquids containing common preservatives like creosote, penta-in-oil, and sodium flouride to treat internal decay, and (3) another liquid for the same purpose which on contact with water changes form into a vapor (gas) becoming a *fumigant* and moving both up-and downward in the wood cells effectively killing decay present and offering lasting protection for extended periods of time.

Paints, varnishes, and stains are used for decorative effects, but they also attord surface protection by retarding moisture changes and thus decreasing checking, warping, and weathering. Such protection is only superficial and they are not wood preservatives.

Fire-retardant chemicals such as ammonium phosphate and sulfate and salts of zinc and boron are used to decrease the flammability of wood. Some fire-retardant formulations may give protection against decay.

Methods of Treating Wood. The methods of preservative treatment are divided into two categories, *pressure* and *nonpressure*. Pressure methods are very effective in the treating of difficult to treat woods such as Douglas-fir, and the processing usually takes much less time if the seasoning does not need to be accomplished in the cylinder than the nonpressure processes, which rely on vacuum, diffusion, and/or temperature differentials to treat the wood. The pressure is conducted in an air-tight cylinder with the preservative liquids being forced into the products, while the nonpressure methods such as dipping, soaking, brushing, and/or spraying are less effective and are not recommended for products requiring trouble-free long service lives or installation into ground exposures. The *vacuum process* can achieve better results and the *thermal process* is capable of treating wood poles for specifications requiring adequate retention and penetration into the wood, sufficient to protect wood for inground installation. The thermal process (formerly the “hot-cold” bath process) accomplishes the required penetration and retention of preservative by first covering the wood with a “hot” liquid at about 112.8°C (23°F), increasing the temperature of the air in the wood cells causing expansion, and then allowing the preservative to either cool down or pumping the hot oil out of the tank and pumping relatively cold preservative into the tank at about 66°C (150°F), which effectively cools the wood cells and the air to contract and form a vacuum pulling the preservative into the cells.

There are several modifications of the pressure process. The full-cell process is used for the treatment of marine piling, which requires high retention of creosote for protection against wood-boring animals. The process is also used commonly in the treatment of lumber with water-borne presevatives. An initial vacuum removes the air from the cell lumens providing maximum penetration and retention into those spaces and the cell walls. Much wood for land use is treated with oil-type preservatives by one of the so-called empty-cell methods, whereby it is possible to increase the depth of penetration obtained with a limited retention of preservative. In the Rueping process, air is first injected to create within the wood a pressure greater than atmospheric. The cylinder is filled with preservative in such a way that the injected air is trapped in the wood. The pressure is then increased to force preservative into the wood. After the pressure is released and the cylinder drained, the compressed air in the wood expands to expel some of the preservative. The recovered preservative is called *kickback.* The Lowry process differs from the Rueping process in that no initial air pressure is applied. In this process, the air normally present in the cells is compressed during the pressure cycle and produces some kickback when pressure is released.

The conditioning of the wood prior to treatment is an important step. Air seasoning, kiln drying, and various processes of cylinder conditioning are employed. The latter include steaming plus vacuum, boiling in oil under vacuum, and vapor drying, in which green wood is surrounded by hot vapors of distillates of coal tar or petroleum.

When oil preservatives are applied by simple soaking methods, the wood should be well-seasoned in order to provide airspaces sans water into which the oil may move. Oil preservatives of low viscosity are preferable. The results attainable vary greatly with the species of wood.

Diffusion methods depend on the diffusion of water-soluble chemicals into the moisture present in green wood. Here again, the species of wood is an important factor, but the results are affected by other factors such as the nature of the chemical, the concentration of the solution, and the duration of the soaking period.

Applications of Preservative Treatment. Preservative treatments are applied to many wood products including poles, crossties, lumber, structural timbers such as crossarms and crossbraces, fence posts, and piling.

Advantages of Preservative Treatment. In addition to the conservation of a natural resource, preservative treatment results in economic savings due to increased service life and reduced maintenance costs. This has been recognized for many years by railroad companies, utility companies, and other large users of wood products. Because of demonstrated savings, practically all crossties and poles are now given a preservative treatment before installation. There has been a gradual increase in the volume of lumber treated annually, due to more widespread knowledge of the need for such treatment when the wood is to be used under conditions favorable to attack by decay or insects. For best performance, it is desirable that all machining operations be completed before treatment.

Strength of Treated Lumber. The effect of a preservative such as creosote or pentachlorophenol, in and of itself, on the strength of treated lumber appears to be negligible. It may be necessary, however, in establishing design stress values, to take into account possible reductions in strength that may result from temperatures or pressures used in the conditioning or treating processes. Results of tests of treated wood show reductions of stress in extreme fiber in bending and compression perpendicular to grain, ranging from a few percent up to 25%, depending on the processes used. Compression parallel to grain is affected less and modulus of elasticity very little. The effect on resistance to horizontal shear can be estimated by inspection for shakes and checks after treatment. Strength reductions for wood poles agreed on in formulating fiber-stress recommendations in American Standard Specifications and Dimensions for Wood Poles, ANSI 05.1, range from 0% to 15% in various species, depending on the conditioning and treating processes. Treating conditions specified by the American Wood-Preservers' Association should never be exceeded. Reductions of strength can be minimized by restricting temperatures, heating periods, and pressures as much as is consistent with obtaining the absorption and penetration required for proper treatment.

Effect of Preservative Treatment on Electrical Resistivity. The electrical resistivity of wood depends on its moisture content to a much greater degree than any other single variable. Ovendry wood is an excellent insulator, but as the wood absorbs moisture, its resistivity decreases rapidly. Wood in normal use, however, where its moisture content may range from about 6% to 14%, is still a good enough insulator for many electrical applications.

When wood has been treated with salts for preservative or fire-retardant purposes, its electrical resistivity may be markedly reduced. The effect of such salt treatment is small when the wood moisture is below about 8% but increases rapidly as the moisture content exceeds about 10%. Treatment with creosote or pentachlorophenol has practically no effect on the resistivity of wood, just as long as no excess water is the treating solution. The resistivity of wood decreases by about a factor of 2 for each increase of 10°C in the temperature and is about half as great for current flow along the grain as across the grain.

4.5.8 American Lumber Standards

Voluntary Product Standard PS 20-99, American softwood Lumber Standard is a voluntary standard of manufacturers, distributors, and users promulgated in cooperation with the U.S. Department of Commerce. It provides for use classifications of (1) yard lumber, (2) structural lumber, and (3) factory and shop lumber. Different grading rules apply to each class. Size standards and generalized grade descriptions are part of PS 20-99, but details of grading rules are left to the organized agencies of the lumber manufacturing industry. The grades and working stresses for structural lumber are referred by PS 20-99 to the authority of ASTM D245, Methods for Establishing Structural Grades of Lumber, or D2018, Recommended for Determining Design Stresses for Load-Sharing Lumber Members.

Standard Commercial Names. Standard commercial names of the most commonly used structural softwood from ASTM D1165, Standard Nomenclature of Domestic Hardwood and Softwoods, are as follows:

- Cedar
  - Alaska cedar
  - Port Orford cedar
  - Western red cedar
- Fir
  - Douglas-fir
  - White fir
- Hemlock
  - Eastern hemlock
  - Western hemlock
- Larch, Western
- Pine
  - Jack pine
  - Lodgepole pine
  - Norway pine
  - Ponderosa pine
  - Southern yellow pine
- Redwood
- Spruce
  - Eastern spruce
  - Engelmann spruce
  - Sitka spruce

Standard Structural Grades. Detailed descriptions of the standard structural grades are published and promulgated in the grading rule books of the organized regional agencies of the lumber manufacturing industry. These are subject to review for compliance with the general requirements of PS 20–99, American softwood Lumber Standard. The principal-use classes of structural lumber are: (1) *joists and planks*, pieces of rectangular cross section 5 to 10 cm (2 to 4 in) thick and 10 cm (4 in) or more wide (nominal dimensions), graded primarily for bending strength edgewise or flatwise, (2) *beams and stringers*, pieces of rectangular cross section 12.7 by 20.3 cm (5 by 8 in) (nominal dimensions) and up, graded for strength in bending when loaded on the narrow face, and (3) *posts and timbers*, pieces of square or nearly square cross section, 12.7 by 12.7 cm (5 by 5 in) (nominal dimensions) and larger, graded primarily for use as posts and columns.

Working Stresses. Working stresses recommended by the lumber industry for their structural grades are found with the detailed grade descriptions in the grading rule books of the organized regional agencies of the industry. A complete listing of all structural grades and their working stresses is found in the *National Design Specification for Wood Construction,* published by ANSI/AF&PAssn. Values for a few typical grades are shown in Table 4-18. Working stresses vary according to the grades and sizes of lumber and their condition with respect to moisture content. Stresses are adjustable also for duration of load and for special conditions such as extreme temperature. Stress increases are provided for "load-sharing/repetitive members" in which the safety of the structure

TABLE 4-18 Design Values for Visually Graded Timbers (5″ × 5″ and larger)*

USE WITH TABLE 4D ADJUSTMENT FACTORS

Species and commercial grade	Size classification	Design values in pounds per square inch (psi)						Grading Rules Agency
		Bending F_b	Tension parallel to grain F_t	Shear parallel to grain F_v	Compression perpendicular to grain F_{c1}	Compression parallel to grain F_c	Modulus of Elasticity E	
DOUGLAS FIR-LARCH								
Dense Select Structural	Beams and Stringers	1900	1100	170	730	1300	1,700,000	West Coast Lumber Inspection Bureau
Select Structural		1600	950	170	625	1100	1,600,000	
Dense No. 1		1550	775	170	730	1100	1,700,000	
No. 1		1350	650	170	625	925	1,600,000	
No. 2		875	425	170	625	600	1,300,000	
Dense Select Structural	Posts and Timbers	1750	1150	170	730	1350	1,700,000	
Select Structural		1500	1000	170	625	1150	1,600,000	
Dense No. 1		1400	950	170	730	1200	1,700,000	
No. 1		1200	825	170	625	1000	1,600,000	
No. 2		750	475	170	625	700	1,300,000	
SOUTHERN PINE		**(Wet Service Conditions)**						
Dense Select Structural	5″ × 5″ & larger	1750	1200	165	440	1100	1,600,000	Southern Pine Inspection Bureau
Select Structural		1500	1000	165	375	950	1,500,000	
No. 1 Dense		1550	1050	165	440	975	1,600,000	
No. 1		1350	900	165	375	825	1,500,000	
No. 2 Dense		975	650	165	440	625	1,300,000	
No. 2		850	550	165	375	525	1,200,000	
Dense Structural 86		2100	1400	165	440	1300	1,600,000	
Dense Structural 72		1750	1200	165	440	1100	1,600,000	
Dense Structural 65		1600	1050	165	440	1000	1,600,000	

Note: 1 lb/in^2 = 6.895 kPa.

*__LUMBER DIMENSIONS__. Tabulated design values are applicable to lumber that will be used under dry conditions such as in most covered structures. For 5″ and thicker lumber, the GREEN dressed sizes shall be permitted to be used (see Table 1A) because design values have been adjusted to compensate for any loss in size by shrinkage which may occur. (Tabulated design values are for normal load duration and dry service conditions, unless specified otherwise. See NDS 4.3 for a comprehensive description of design value adjustment factors.)

Source: Compiled from *National Design Specification for Wood Construction* ANSI/AF$P ASSN.-1 depends upon the strength of the assemblage of members rather than upon the lowest strength value for any single member. These stress modifications are described in ASTM standards.

Allowable working stresses for the structural grades of lumber are also a part of certain use specifications, such as the Minimum Property Standards of the Federal Housing Administration, the *American Railway Engineering Association Manual*, and various local or regional building codes. These allowable values may or may not coincide with the lumber industry stress recommendations for the same species and grade.

Wood-Base Panel Materials. Included in this category are insulating board, hardboard, particle board, waferboard, plywood, and several engineered-wood products. Plywood, normally fabricated by bonding an odd number of layers of veneers together with the grain direction in adjacent piles at right angles to each other, is more dimensionally stable and more uniform in strength in the plane of the sheet than wood. Qualities of glue line and veneer permitted are set by the various commercial standards for plywood and determine the grades under which plywood is sold. In general, glue-line quality determines whether plywood is classed as being suitable for interior or exterior use.

U.S. Product Standard PS1-83 covers the basic specifications for the manufacture of construction plywood. Decorative hardwood plywood is described by U.S. Product Standard PS51(5.2). Plywood manufactured according to this standard will carry a grade trademark of a qualified testing agency. Testing is conducted under PS2-92 and ASTM D1037.

Insulation boards and hardboards are panel products made by reducing wood substance to particles or fiber and reconstituting the fiber into stiff panels 1.2 by 2.4 m (4 by 8 ft) in area or larger. Insulation board is of either interior or water-resistant quality, and is usually manufactured for use where combinations of thermal and sound-insulating properties and stiffness and strength are desired. Hardboard with a density of 50 lb/ft^3 or more is used in many applications where a relatively thin, hard, uniform panel material is required. Of great importance in the electrical field are special high-density hardboard products expressly manufactured with high dielectric properties. Medium-density fiberboards have a density between that of insulation board and hardboard.

Particle boards are panel products made by gluing small pieces of wood in a form such as flakes (this product is called waferboard) and shavings into relatively thick, rigid panels. Thermosetting resins are used to provide bonds of either interior or water-resistant quality. Standards ASTM C208 and ANSI A208.1 and PS58 govern minimum qualities of regular insulation board, particle board, and hardboard. The important physical and strength properties of various board products are indicated by Table 4-19.

There are many engineered-wood products now available for structural use. In addition to plywood, which has already been discussed, they include glued-laminated timber, wood I-joists,oriented-strand board (OSB), laminated-veneer lumber (LVL), parallel-strand lumber (PSL), and laminated-strand lumber (LSL). All these products require structural engineering input in their design, manufacture, and end use. There are many advantages in using these products compared with solid wood.

4.5.9 Wood Poles and Crossarms

Douglas-fir (Coast) and Southern pine as a group of four species, loblolly, longleaf, shortleaf, and slash pine, are the most commonly used in the United States to support electric supply and communication lines, cables, and equipment. In the Northeast part of the country, there are still some chestnut and also Northern white cedar poles in service, but these species are no longer available for purchase. Western red cedar, lodgepole and ponderosa pines, and Western larch are used to a limited extent in the Western states. Other species are shown in the ANSI 05.1 Specification size/class tables, but many of those are not provided for orders due to lack of supply or lack of fabricating and treating sources.

4.5.10 Standards for Wood Poles

The ANSI specifications for wood poles serve as a basis for purchasing and use. The ANSI specifications cover fiber stresses, dimensions, defect limitations, and manufacturing requirements. The fiber stresses approved by the ANSI and contained in its Standard 05.1 are as shown in Tables 4-20, 4-21, 4-22, and 4-23. These tables cover all species of poles normally used in communication and electrical power construction.

TABLE 4-19 Strength and Mechanical Properties of Selected Wood-Based Composite Products*

Material	Density, lb/ft	Specific gravity	Modulus of rupture, lb/in²	Modulus of elasticity (bending), 1,000 lb/in²	Tensile strength parallel to surface, lb/in²	Tensile strength perpendicular to surface, lb/in²	Compression strength parallel to surface, lb/in²	24-h water absorption, % by volume	24-h water absorption, % by weight	Thickness swelling, 24-h soak, %	Maximum linear expansion,† %	Thermal conductivity Btu/(ft²)(h)(°F) (in thickness)
Fibrous-felted boards:												
1. Structural insulating board	10–26	0.16–0.42	200–800	25–125	200–500	10–25	—	1–10	—	—	0.5	0.27–0.45
2. Hardboard												
a. Tempered	60–80	0.96–1.28	6,000	800–1,000	3,000	130	4,200	—	10–30	9–25	0.4	1.10–1.50
b. Standard	50–80	0.80–1.28	4,500	400–800	2,200	90	1,800	—	15–40	10–30	0.6	0.8–1.40
Particle boards:												
1. High density, grade I-H-2	50–70	0.8–1.12	3,000	350	—	130	—	—	—	—	—	0.4–1.0
2. Medium density, grade I-M-2	37–50	0.6–0.8	2,100	325	—	60	—	—	—	—	—	1.1–1.5
3. Waferboard	37–42	0.6–0.67	2.500	450	—	50	—	—	—	20–25	0.2	
Veneered boards:												
1. Plywood with Douglas fir or southern pine faces, dry use:												
a. Grade stress level S-1	30–37	0.46–0.59	2,000	1,800	1,400	120	1,060					
b. Grade stress level S-2	30–37	0.46–0.59	1,650	1,800	1,200	120	990					
2. Paralam	30–37	0.46–0.59	2,900	2,000	2,400		2,900					

Note: 1 lb/ft = 1.88 kg/m; 1 lb/in² = 6.895 kPa.

*The data presented are general round-figure values, accumulated from numerous sources; for more exact figures on a specific product, individual manufacturers should be consulted or actual tests made. Values are for general laboratory conditions of temperature and relative humidity.

†Expansion resulting from a change in moisture content from equilibrium at 50% relative humidity to equilibrium at 90% relative humidity.

TABLE 4-20 Dimensions of Northern White Cedar and Engelmann Spruce Poles

	Class		1	2	3	4	5	6	7	9	10
	Minimum circumference at top, in		27	25	23	21	19	17	15	15	12
	Length of pole, ft	Ground-line distance from butt,* ft	Minimum circumference at 6 ft from butt, in								
Northern white cedar poles (based on a fiber stress of 4,000 lb/in^2)	20	4	38.0	35.5	33.0	30.5	28.0	**26.0**	**24.0**	**22.0**	**17.5**
	25	5	42.0	39.5	36.5	34.0	31.5	**29.0**	**27.0**	**24.0**	**19.5**
	30	5½	45.5	**43.0**	**40.0**	**37.0**	**34.5**	**32.0**	**29.5**	**26.0**	
	35	6	49.0	**46.0**	**42.5**	**39.5**	**37.0**	**34.0**	**31.5**		
	40	6	**51.5**	**48.5**	**45.0**	**42.0**	**39.0**	**36.0**			
	45	6½	**54.5**	**51.0**	**47.5**	**44.0**	**41.0**				
	50	7	**57.0**	**53.5**	**49.5**	46.0	43.0				
	55	7½	**59.0**	**55.5**	**51.5**	48.0	44.5				
	60	8	**61.0**	**57.5**	**53.5**	50.0					
Engelmann spruce poles (based on a fiber stress of 5,600 lb/in^2)	20	4	34.5	32.0	30.0	28.0	25.5	**23.5**	**22.0**	**19.0**	**15.0**
	25	5	38.0	35.5	33.0	30.5	28.5	**26.0**	**24.5**	**21.0**	**16.5**
	30	5½	41.0	**38.5**	**35.0**	**33.0**	**30.5**	**28.5**	**26.5**	**22.5**	
	35	6	43.5	**41.0**	**38.0**	**35.5**	**32.5**	**30.5**	**28.0**		
	40	6	**46.0**	**43.5**	**40.5**	**37.5**	**34.5**	**32.0**			
	45	6½	**48.5**	**45.5**	**42.5**	**39.5**	**36.5**				
	50	7	**50.5**	**47.5**	**44.5**	41.0	38.0				
	55	7½	**52.5**	**49.5**	**46.0**	42.5	39.5				
	60	8	**54.5**	**51.0**	**47.5**	44.0					
	65	8½	**56.0**	**52.5**	**49.0**	45.5					
	70	9	**57.5**	**54.0**	**50.5**	47.0					
	75	9½	**59.5**	**55.5**	**52.0**						
	80	10	**61.0**	**57.0**	53.5						
	85	10½	**62.5**	**58.5**	54.5						
	90	11	**63.5**	**60.0**	56.0						
	95	11	**65.0**	**61.0**							
	100	11	**66.0**	**62.0**							

Notes: Classes and lengths for which circumferences at 6 ft (1.83 m) from the butt are listed in boldface type are the preferred standard sizes. Those shown in light type are included for engineering purposes only.

1 in = 2.54 cm; 1 ft = 0.3048 m; 1 lb/in^2 = 6.895 kPa.

*The figures in this column are intended for use only when a definition of ground line is necessary in order to apply requirements relating to scars, straightness, etc.

Pole Dimensions. The circumference at "6 ft (1.8 m) from butt" in ANSI Standard 05.1 is based on the following principles:

a. The classes from the lowest to the highest were arranged in approximate geometric progression, the increments in breaking load between classes being about 25%.

b. The dimensions were specified in terms of circumference in inches at the top and circumference in inches at 6 ft from the butt for poles of the respective classes and lengths, except for three classes having no requirement for butt circumference.

c. All poles of the same class and length were to have, when new, approximately equal strength or, in more precise terms, equal moments of resistance at the ground line.

d. All poles of different lengths within the same class were of sizes suitable to withstand approximately the same breaking load, on the assumption that the load is applied 0.6 m (2 ft) from the top and that the break (failure) would occur at the ground line.

TABLE 4-21 Western Red Cedar, Ponderosa Pine, and Poles of Other Species

	Class		1	2	3	4	5	6	7	9	10
	Minimum circumference at top, in		27	25	23	21	19	17	15	15	12
	Length of pole, ft	Ground-line distance from butt,* ft	Minimum circumference at 6 ft from butt, in								
Western red cedar and ponderosa pine poles (based on a fiber stress of 6,000 lb/in²)	20	4	33.5	31.5	29.5	27.0	25.0	**23.0**	**21.5**	**18.5**	**15.0**
	25	5	37.0	34.5	32.5	30.0	28.0	**25.5**	**24.0**	**20.5**	**16.5**
	30	$5\frac{1}{2}$	40.0	**37.5**	**35.0**	**32.5**	**30.0**	**28.0**	**26.0**	**22.0**	
	35	6	42.5	**40.0**	**37.5**	**34.5**	**32.0**	**30.0**	**27.5**		
	40	6	**45.0**	**42.5**	**39.5**	**36.5**	**34.0**	**31.5**			
	45	$6\frac{1}{2}$	**47.5**	**44.5**	**41.5**	**38.5**	**36.0**	33.0			
	50	7	**49.5**	**46.5**	**43.5**	40.0	37.5				
	55	$7\frac{1}{2}$	**51.5**	**48.5**	**45.0**	42.0					
	60	8	**53.5**	**50.0**	**46.5**	43.5					
	65	$8\frac{1}{2}$	**55.0**	**51.5**	**48.0**	45.0					
	70	9	**56.5**	**53.0**	**49.5**	46.0					
	75	$9\frac{1}{2}$	**58.0**	**54.5**	**51.0**						
	80	10	**59.5**	**56.0**	52.0						
	85	$10\frac{1}{2}$	**61.0**	**57.0**	53.5						
	90	11	**62.5**	**58.5**	54.5						
	95	11	**63.5**	**59.5**							
	100	11	**65.0**	**61.0**							
	105	12	**66.0**	**62.0**							
	110	12	**67.5**	**63.0**							
	115	12	**68.5**	**64.0**							
	120	12	**69.5**	**65.0**							
	125	12	**70.5**	**66.0**							

	Class		1	2	3	4	5	6	7	9	10
	Minimum circumference at top, in		27	25	23	21	19	17	15	15	12
	Length of pole, ft	Ground-line distance from butt,* ft	Minimum circumference at 6 ft from butt, in								
Jack pine, lodgepole pine, red pine, redwood, Sitka spruce, Western fir, and white spruce poles (based on a fiber stress of 6,600 lb/in²)	20	4	32.5	30.5	28.5	26.5	24.5	**22.5**	**21.0**	**18.0**	**14.5**
	25	5	36.0	33.5	31.0	29.0	27.0	**25.0**	**23.0**	**20.0**	**15.5**
	30	$5\frac{1}{2}$	39.0	**36.5**	**34.0**	**31.5**	**29.0**	**27.0**	**25.0**	**21.0**	
	35	6	41.5	**38.5**	**36.0**	**33.5**	**31.0**	**28.5**	**26.5**		
	40	6	**44.0**	**41.0**	**38.0**	**35.5**	**33.0**	**30.5**			
	45	$6\frac{1}{2}$	**46.0**	**43.0**	**40.0**	**37.0**	**34.5**	32.0			
	50	7	**48.0**	**45.0**	**42.0**	39.0	36.0				
	55	$7\frac{1}{2}$	**49.5**	**46.5**	**43.5**	40.5					
	60	8	**51.5**	**48.0**	**45.0**	42.0					
	65	$8\frac{1}{2}$	**53.0**	**49.5**	**46.0**	43.0					
	70	9	**54.5**	**51.0**	**47.5**	44.5					
	75	$9\frac{1}{2}$	**56.0**	**52.5**	**49.0**						
	80	10	**57.5**	**54.0**	50.5						
	85	$10\frac{1}{2}$	**58.5**	**55.0**	51.5						
	90	11	**60.0**	**56.5**	52.5						
	95	11	**61.5**	**57.5**							
	100	11	**62.5**	**58.5**							
	105	12	**63.5**	**60.0**							
	110	12	**65.0**	**61.0**							
	115	12	**66.0**	**62.0**							
	120	12	**67.0**	**63.0**							
	125	12	**68.0**	**64.0**							

TABLE 4-22 Alaska Yellow Cedar, Western Hemlock, Douglas Fir, and Southern Pine Poles

	Class		1	2	3	4	5	6	7	9	10
	Minimum circumference at top, in		27	25	23	21	19	17	15	15	12
	Length of pole, ft	Ground-line distance from butt,* ft	Minimum circumference at 6 ft from butt, in								
Alaska yellow cedar and Western hemlock poles (based on a fiber stress of 7,400 lb/in^2)	20	4	31.5	29.5	27.5	25.5	23.5	**22.0**	**20.0**	**17.5**	**14.0**
	25	5	34.5	32.5	30.0	28.0	26.0	**24.0**	**22.0**	**19.5**	**15.0**
	30	5½	37.5	**35.0**	**32.5**	**30.0**	**28.0**	**26.0**	**24.0**	**20.5**	
	35	6	40.0	**37.5**	**35.0**	**32.0**	**30.0**	**27.5**	**25.5**		
	40	6	**42.0**	**39.5**	**37.0**	**34.0**	**31.5**	**29.0**			
	45	6½	**44.0**	**41.5**	**38.5**	**36.0**	**33.0**	30.5			
	50	7	**46.0**	**43.0**	**40.0**	37.5	34.5				
	55	7½	**47.5**	**44.5**	**41.5**	39.0					
	60	8	**49.5**	**46.0**	**43.0**	40.0					
	65	8½	**51.0**	**47.5**	**44.5**	41.5					
	70	9	**52.5**	**49.0**	**46.0**	42.5					
	75	9½	**54.0**	**50.5**	**47.0**						
	80	10	**55.0**	**51.5**	48.5						
	85	10½	**56.5**	**53.0**	49.5						

	Class		1	2	3	4	5	6	7	9	10
	Minimum circumference at top, in		27	25	23	21	19	17	15	15	12
	Length of pole, ft	Ground-line distance from butt,* ft	Minimum circumference at 6 ft from butt, in								
Douglas-fir and Southern pine poles (based on a fiber stress of 8.000 lb/in^2	20	4	31.0	29.0	27.0	25.0	23.0	**21.0**	**19.5**	**17.5**	**14.0**
	25	5	33.5	31.5	29.5	27.5	25.5	**23.0**	**21.5**	**19.5**	**15.0**
	30	5½	36.5	**34.0**	**32.0**	**29.5**	**27.5**	**25.0**	**23.5**	**20.5**	
	35	6	39.0	**36.5**	**34.0**	**31.5**	**29.0**	**27.0**	**25.0**		
	40	6	**41.0**	**38.5**	**36.0**	**33.5**	**31.0**	**28.5**			
	45	6½	**43.0**	**40.5**	**37.5**	**35.0**	**32.5**	30.0			
	50	7	**45.0**	**42.0**	**39.0**	36.5	34.0				
	55	7½	**46.5**	**43.5**	**40.5**	38.0					
	60	8	**48.0**	**45.0**	**42.0**	39.0					
	65	8½	**49.5**	**46.5**	**43.5**	40.5					
	70	9	**51.0**	**48.0**	**45.0**	41.5					
	75	9½	**52.5**	**49.0**	**46.0**						
	80	10	**54.0**	**50.5**	47.0						
	85	10½	**55.0**	**51.5**	48.0						
	90	11	**56.0**	**53.0**	49.0						
	95	11	**57.0**	**54.0**							
	100	11	**58.5**	**55.0**							
	105	12	**59.5**	**56.0**							
	110	12	**60.5**	**57.0**							
	115	12	**61.5**	**58.0**							
	120	12	**62.5**	**59.0**							
	125	12	**63.5**	**59.5**							

Notes: Classes and lengths for which circumferences at 6 ft (1.83 cm) from the butt are listed in boldface type are the preferred standard sizes. Those shown in light type are included for engineering purposes only.

1 in = 2.54 cm; 1 ft = 0.3048 m; 1 lb/in^2 = 6.895 kPa.

*The figures in this column are intended for use only when a definition of ground line is necessary in order to apply requirements relating to scars, straightness, etc.

TABLE 4-23 Western Larch Poles

	Class		1	2	3	4	5	6	7	9	10
	Minimum circumference at top, in		27	25	23	21	19	17	15	15	12
	Length of pole, ft	Ground-line distance from butt,* ft	Minimum circumference at 6 ft from butt, in								
	90	11	**57.5**	**54.0**	50.5						
	95	11	**58.5**	**55.0**							
	100	11	**60.0**	**56.0**							
	105	12	**61.0**	**57.0**							
	110	12	**62.0**	**58.0**							
	115	12	**63.0**	**59.0**							
	120	12	**64.0**	**60.0**							
	125	12	**65.0**	**61.0**							
Western larch poles (based on a fiber stress of 8,400 lb/in^2)	20	4	30.0	28.5	26.5	24.5	22.5	**21.0**	**19.0**	**17.0**	**13.5**
	25	5	33.0	31.0	29.0	26.5	24.5	**23.0**	**21.0**	**18.5**	**14.5**
	30	5$^1/_2$	35.5	**33.5**	**31.0**	**29.0**	**26.5**	**24.5**	**23.0**	**19.5**	
	35	6	38.0	**35.5**	**33.0**	**31.0**	**28.5**	**26.5**	**24.5**		
	40	6	**40.0**	**37.5**	**35.0**	**32.5**	**30.0**	**28.0**			
	45	6$^1/_2$	**42.0**	**39.5**	**37.0**	**34.0**	**31.5**	29.0			
	50	7	**44.0**	**41.0**	**38.5**	35.5	33.0				
	55	7$^1/_2$	**45.5**	**42.5**	**40.0**	37.0					
	60	8	**47.0**	**44.0**	**41.0**	38.5					
	65	8$^1/_2$	**48.5**	**46.0**	**42.5**	39.5					
	70	9	**50.0**	**47.0**	**44.0**	41.0					
	75	9$^1/_2$	**51.5**	**48.0**	**45.0**						
	80	10	**52.5**	**49.5**	46.0						
	85	10$^1/_2$	**54.0**	**50.5**	47.0						
	90	11	**55.0**	**51.5**	48.5						
	95	11	**56.5**	**53.0**							
	100	11	**57.5**	**54.0**							
	105	12	**58.5**	**55.0**							
	110	12	**59.5**	**56.0**							
	115	12	**60.5**	**57.0**							
	120	12	**61.5**	**58.0**							
	125	12	**62.5**	**58.5**							

Notes: Classes and lengths for which circumferences at 6 ft (1.83 m), from the butt are listed in boldface type are the preferred standard sizes. Those shown in light type are included for engineering purposes only.

1 in = 2.54 cm; 1 ft = 0.3048 m; 1 lb/in^2 = 6.895 kPa.

*The figures in this column are intended for use only when a definition of ground line is necessary in order to apply requirements relating to scars, straightness, etc.

The breaking loads referred to in (d) above for the classes for which 6 ft (1.8 m) from butt circumferences are given are as follows: Class 1, 4,500 lb; Class 2, 3,700 lb; Class 3, 3,000 lb; Class 4, 2,400 lb; Class 5, 1,900 lb; Class 6, 1,500 lb; Class 7, 1,200 lb; Class 9, 740 lb; Class 10,370 lb.

Minimum top circumferences and minimum circumferences at 6 ft (1.8 m) from butt are given in Tables 4-20, 4-21, 4-22, and 4-23.

Length. Poles under 50 ft (15.2 m) in length should not be more than 3 in shorter or 6 in longer than nominal length. Poles 50 ft (15.2 m) or over in length should not be more than 6 in shorter or 12 in longer than nominal length. Length should be measured between the extreme ends of the pole.

Circumference. The minimum circumference at 6 ft (1.8 m) from the butt and at the top, for each length and class of pole, is listed in the tables of dimensions. The circumference at 6 ft (1.8 m) from the butt of poles should be not more than 7 in (17.8 cm) or 20% larger than the specified minimum, whichever is greater. The top dimensional requirement should apply at a point corresponding to the minimum length permitted for the pole.

Classification. The true circumference class should be determined as follows: Measure the circumference at 6 ft (1.8 m) from the butt. This dimension will determine the true class of the pole, provided that its top (measured at the minimum length point) is large enough. Otherwise, the circumference at the top will determine the true class, provided that the circumference at 6 ft (1.8 m) from the butt does not exceed the specified minimum by more than 7 in (17.8 cm) or 20%, whichever is greater. The preceding information relating to the pole standards approved by the ANSI does not constitute the complete standards. For further information, consult the standards, which may be obtained at a nominal charge.

Machine shaving of poles has increased as a practice of producers. Some producers also turn the pole down in the process, thereby obtaining a straighter pole with a specific taper. The machine-processed poles season more rapidly, which is particularly important with species like Southern pine which is susceptible to fungus attack before treatment. Machine shaving makes for easier detection of defects and provides a pole of improved appearance. If poles having normally thin sapwood (such as Western red cedar and larch) are to be full-length when treated with preservative, it is undesirable to reduce the thickness of the sapwood more than necessary to obtain a dressed pole.

Inspection. Poles are inspected prior to treatment for physical defects and decay and after treatment for penetration and retention of preservative and for cleanliness. Inspection is most effective when made at vendors' plants, because defects that may be hidden by preservative are detected and freight is saved on rejects. Commercial inspection agencies are available at most producing locations, and it is normally economical to utilize their services. Quantity users may have their own trained inspectors. *Crossarm inspection* is important because safety of linemen is a consideration in addition to quality of timber. As with poles, inspection should be made before treatment for defects and after treatment for penetration, retention, and cleanliness. Inspection should be done by qualified timber specialists.

Conductivity. Conductivity is of concern to many electric-utility companies. Pole resistance varies greatly with moisture content. Dry wood of all species exhibits high resistance. Surface absorption of rainwater by untreated wood may vary the resistance over a wide range. Full-length treated poles thoroughly dried before treatment generally show only moderate reduction in resistance following a rain. A rough correlation between resistance and moisture will show that 500,000 Ω over a 20-ft length of pole between contacts driven $3\frac{1}{2}$ in deep corresponds to a moisture content of about 25%. Other average points on the curve band are 50,000,000 Ω 15% moisture and 20,000 Ω 40% moisture.

Depth of Pole Setting. The values in the column headed "Ground-line distance from butt" in Tables 4-20 to 4-23 may be accepted as a guide for a satisfactory depth of pole settings in ordinary firm soil. In marshy soil and at unguyed angles in lines, setting depths should be increased 1 to 2 ft (0.3 to 0.6 m). In rock, the indicated settings may be reduced one-half for that part of the pole set in rock. Rock backfill in ordinary earth locations is not considered as set in rock.

Pole Stubbing. Pole stubbing frequently can be employed to effect substantial money savings. An otherwise good pole that is decayed at or below the ground line is fastened securely to a new preservative-treated stub set in the ground alongside it. The major part of the savings resulted from avoidance of transferring wires and equipment.

Salvaging. Poles removed for any reason can frequently be salvaged for future use. Users of large quantities can economically do this. The following operations must be employed: if defective cut off top, butt and retreat, remove old hardware, peel off saprot, reframe, and retreat as required.

Kinds of Timber for Crossarms. Two kinds of timber are in general use for crossarms, Douglas-fir and Southern pine. All pine crossarms are treated with creosote or pentachlorophenol. Crossarms must be treated to achieve long service life.

Most Douglas-fir arms used for communication and power-distribution lines are manufactured from timber selected for the purpose. Dense and close-grain lumber is used. Publication 17 of the West Coast Lumber Inspection Bureau sets forth grading and dressing requirements.

TABLE 4-24 Bending Load and Crushing Strength of Crossarms

Species	Rings per in	Percent Summer wood	Percent Sap wood	Percent Moisture	Density (dry)	Max. bending load, lb	Maximum crushing strength lb/sq in
Douglas fir	20	40	0	11.5	0.48	7,590	7,080
Longleaf pine (50% heart)	18	44	55	13.4	0.54	8,984	5,425
Longleaf pine (75% heart)	19	53	32	13.5	0.63	10,180	8,950
Longleaf pine (100% heart)	16	44	1	12.8	0.63	9,782	8,940
Shortleaf pine	11	46	79	13.3	0.52	9,260	7,300
Shortleaf pine, creosoted	11	49	. . .	. . .	. . .	7,649	5,770
White cedar	12	45	2	14.3	0.36	5,200	4,700

Note: 1 in = 2.54 cm; 1 lb = 0.4536 kg; 1 lb/in^2 = 6.895 kPa.

There is no grade of Southern pine timber designated as crossarm stock, and crossarm users depend on the limitations set forth in their specifications to obtain a satisfactory quality of product. Pine arms are usually small boxed heart timbers.

Large transmission-line structures are treated round poles and over the past 20 years, laminated towers, arms and braces have more frequently been selected.

Crossarm Specifications. The most widely used specifications for power-line crossarms have been promulgated by ANSI 05.3 as "Solid Sawn-Wood Crossarms and Braces—Specifications and Dimensions" and by the Transmission and Distribution Committee of the Edison Electric Institute. For fir crossarms: Specification TD-90, which combines both dense and close-grain grades; Specification TD-92, Heavy-Duty Douglas-Fir Crossarms; Specification TD-93, Heavy-Duty Douglas-Fir Braces. For pine crossarms: Specification TD-91, Dense Southern Pine Crossarms Preservative Treated. Widely used specifications for communication crossarms are American Telephone and Telegraph Co. Specification AT-7298. Crossarms. Sawn crossarms are treated with preservatives by pressure processes following AWPA Standard C25 laminated structures must confirm with the standards of AITC (American Institute of Timber Construction).

Strength of Crossarms. The most reliable source of information on the strength comes from tests made under conditions to simulate crossarms in service. Some tests have been made, and others are under consideration. Theoretical considerations, treating a crossarm as a beam, are valuable if those factors which control the actual strength are taken into account. Tests made several years ago on 84 six-pin, 3$3\frac{1}{4}$ by 4$3\frac{1}{4}$-in by 6-ft communication crossarms, with a uniformly distributed vertical load, gave average results shown in Table 4-24 (*U.S. Forest Service Circ.* 204, by T. R. C. Wilson). The maximum bending load is the total distributed vertical load. The maximum crushing strength is under compression parallel to the grain. Methods of tests are covered by the appropriate ASTM standards.

BIBLIOGRAPHY

AITC (American Institute of Timber Construction), *Timber Construction Manual*, 5th ed., John Wiley & Sons, 2005.

American Society of Testing Materials (ASTM), West Conshohocken, PA. http://www.astm.org

ANSI (American National Standards Institute), *National Design Specification for Wood Construction* 2001 Edition, ANSI/AF&PA NDS-2001 (revised standard), approval date: November 30, 2001.

APA, The Engineered Wood Products Association, 2004.

ASTM, American Society for Testing and Materials, 2004.

AWPA, American Wood-Preservers' Association, 2004.

Brown, Panshin and Forsaith, *Textbook of Wood Technology*, vol. I, McGraw-Hill, New York, 1949.

Brown, Panshin and Forsaith, *Textbook of Wood Technology* vol. II, McGraw-Hill, New York, 1952.

Christopher, V. F., *Handbook of Electrical Tables and Design*, McGraw-Hill, New York, 1998.

Copper Development Association, New York. http://www.copper.org

Fahery, K. F. and Williamson, T. G., *Wood Engineering & Construction Handbook*, McGraw-Hill, New York, 1989.

Hunt and Garrett, *Wood Preservation*, McGraw-Hill, New York, 1953.

McPartland, Brian J., *National Electrical Code®*, McGraw-Hill, New York, 2005.

National Electrical Manufacturers Association, Rosslyn, VA. http://www.nema.org

National Design Specification—Design Values for Wood Construction NDS 2001 Edition, Supplement.

National Design Specification—Manual for Engineered Wood Construction 2001 Edition; Supplements-ASD for "Structural Lumber, Structural Glued Laminated Timber, Timber Poles and Piles, Wood Structural Panels and Wood Structural Panel Shear Wall and Diaphragm."

National Design Specification—Manual for Engineered Wood Construction 2001 Edition; Guidelines-ASD for "Wood I-Joists, Structural Composite Lumber, Metal Plate Connected Wood Trusses and Pre-Engineered Metal connectors."

National Design Specification—Manual for Engineered Wood Construction 2001 Edition; Supplement—ASD/LRFD for "Special Design Provisions for Wind and Seismic."

National Design Specification—Manual for Engineered Wood Construction 2001 Edition; ASD.

Nicholas, D. D., *Wood Deterioration and Its Prevention by Preservative Treatments*, Syracuse University Press, Syracuse, N. Y., 1973.

Smulski, S., *Engineered Wood Products: A Guide for Specifiers, Designers and Users*, PFS Research Foundation, 1997.

The Aluminum Association, Arlington, VA. http://www.aluminum.org

USDA, *The Wood Handbook: Wood as an Engineering Material*, Handbook no. 72, Forest Products Laboratory, 1999.

USDA, *Dry Kiln Operator's Manual*, SIMPSON, W.T., ed., Handbook no. 188, 1991.

USDA, *Utilization of Hardwoods Growing on Southern Pine Sites*, KOCH, Handbook no. 605, vols. I, II, and III, 1985.

USDA, *Utilization of Southern Pines*, KOCH, Handbook no. 420, vols. I and II, 1972.

Wangaard, *Mechanical Properties of Wood*, John Wiley & Sons, 1950.

Zabel, R. A. and Morrell, J., *Wood Microbiology, Decay and Its Prevention*, Academic Press, New York, 1992.

SECTION 5
GENERATION

Stephen O. Dean
President, Fusion Power Associates

George H. Miley
Department of Nuclear Engineering, University of Illinois

CONTENTS

5.1 FOSSIL-FUELED PLANTS

5.1.1 Introduction

America—and much of the world—is becoming increasingly electrified. In 2005, more than half of the electricity generated in the United States came from coal. For the foreseeable future, coal will continue to be the dominant fuel used for electric power production. The low cost and abundance of coal is one of the primary reasons why consumers in the United States benefit from some of the lowest electricity rates of any free-market economy.

The key challenge to keeping coal viable as a generation fuel is to remove the environmental objections to the use of coal in power plants. New technologies are being developed that could virtually eliminate the sulfur, nitrogen, and mercury pollutants released when coal is burned. It may also be possible to capture greenhouse gases that are emitted from coal-fired power plants and prevent them from contributing to global warming concerns.

Research is also underway to increase the fuel efficiency of coal-fueled power plants. Today's plants convert only one-third of coal's energy potential to electricity. New technologies could nearly double efficiency levels in the next 10 to 15 years.

Natural gas is the fastest growing fuel for electricity generation. More than 90% of the power plants to be built in the next 20 years will likely be fueled by natural gas. Natural gas is also likely to be a primary fuel for distributed power generators—mini-power plants that could be sited close to where the electricity is needed.

Natural gas-powered fuel cells are also being developed for future distributed generation applications. Fuel cells use hydrogen that can be extracted from natural gas, or perhaps in the future from biomass or coal.

5.1.2 Thermodynamic Cycles

Rankine Cycle. The cornerstone of the modern steam power plant is a modification of the Carnot cycle proposed by W. J. M. Rankine, a distinguished Scottish engineering professor of thermodynamics and applied mechanics. The temperature-entropy and enthalpy-entropy diagrams of Fig. 5-1 illustrate the state changes for the Rankine cycle. With the exception that compression terminates (state *a*) at boiling pressure rather than the boiling temperature (state *á*), the cycle resembles a Carnot

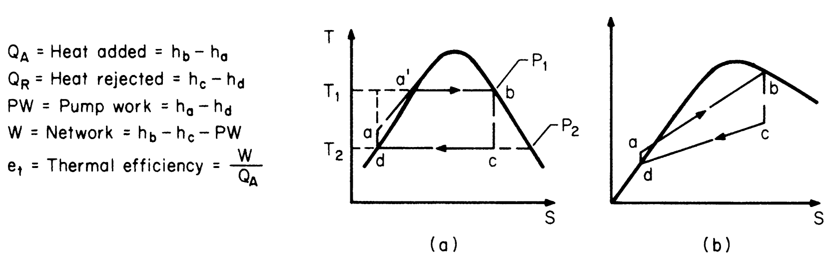


FIGURE 5-1 Simple Rankine cycle (without superheat): (*a*) temperature-entropy; (*b*) enthalpy-entropy (Mollier).

cycle. The triangle bounded by *a-á* and the line connecting to the temperature-entropy curve in Fig. 5-1*a* signify the loss of cycle work because of the irreversible heating of the liquid from state *a* to saturated liquid. The lower pressure at state *a*, compared to *á*, makes possible a much smaller work of compression between *d-a*. For operating plants, it amounts to 1% or less of the turbine output.

This modification eliminates the two-phase vapor compression process, reduces compression work to a negligible amount, and makes the Rankine cycle less sensitive than the Carnot cycle to the irreversibilities bound to occur in an actual plant. As a result, when compared with a Carnot cycle operating between the same temperature limits and with realistic component efficiencies, the Rankine cycle has a larger network output per unit mass of fluid circulated, smaller size, and lower cost of equipment. In addition, because of its relative insensitivity to irreversibilities, its operating plant thermal efficiencies will exceed those of the Carnot cycle.

Regenerative Rankine Cycle. Refinements in component design soon brought power plants based on the Rankine cycle to their peak thermal efficiencies, with further increases realized by modifying the basic cycle. This occurred through increasing the temperature of saturated steam supplied to the turbine, by increasing the turbine inlet temperature through constant-pressure superheat, by reducing the sink temperature, and by reheating the working vapor after partial expansion followed by continued expansion to the final sink temperature. In practice, all of these are employed with yet another important modification. The irreversibility associated with the heating of the compressed liquid to saturation by a finite temperature difference is the primary thermodynamic cause of lower thermal efficiency for the Rankine cycle. The regenerative cycle attempts to eliminate this irreversibility by using as heat sources other parts of the cycle with temperatures slightly above that of the compressed liquid being heated.

This procedure of transferring heat from one part of a cycle to another in order to eliminate or reduce external irreversibilities is called "regenerative heating," which is basic to all regenerative cycles.

The scheme shown in Fig. 5-2 is a practical approach to regeneration. Extraction or "bleeding" of steam at state *c* for use in the "open" heater avoids excessive cooling of the vapor during turbine expansion; in the heater, liquid from the condenser increases in temperature by ΔT. (Regenerative cycle heaters are called "open" or "closed" depending on whether hot and cold fluids are mixed directly to share energy or kept separate with energy exchange occurring by the use of metal coils.)

The extraction and heating substitute the finite temperature difference ΔT for the infinitesimal dT used in the theoretical regeneration process. This substitution, while failing to realize the full potential of regeneration, halves the temperature difference through which the condensate must be heated in the basic Rankine cycle. Additional extractions and heaters permit a closer approximation to the maximum efficiency of the idealized regenerative cycle, with further improvement over the simple Rankine cycle shown in Fig. 5-1.

Reducing the temperature difference between the liquid entering the boiler and that of the saturated fluid increases the cycle thermal efficiency. The price paid is a decrease in net work produced per pound of vapor entering the turbine and an increase in the size, complexity, and initial cost of the plant. Additional improvements in cycle performance may be realized by continuing to accept the consequences of increasing the number of feedwater heating stages. Balancing cycle thermal efficiency against plant size, complexity, and cost for production of power at minimum cost determines the optimum number of heaters.

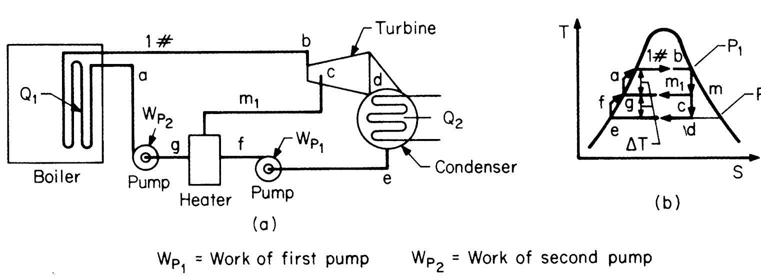


FIGURE 5-2 Single extraction regenerative cycle: (*a*) flow diagram; (*b*) temperature-entropy diagram.

Reheat Cycle. The use of *superheat* offers a simple way to improve the thermal efficiency of the basic Rankine cycle and reduce vapor moisture content to acceptable levels in the low-pressure stages of the turbine. But with continued increase of higher temperatures and pressures to achieve better cycle efficiency, in some situations available superheat temperatures are insufficient to prevent excessive moisture from forming in the low-pressure turbine stages.

The solution to this problem is to interrupt the expansion process, remove the vapor for *reheat* at constant pressure, and return it to the turbine for continued expansion to condenser pressure. The thermodynamic cycle using this modification of the Rankine cycle is called the "reheat cycle." Reheating may be carried out in a section of the boiler supplying primary steam, in a separately fired heat exchanger, or in a steam-to-steam heat exchanger. Most present-day utility units combine superheater and reheater in the same boiler.

Usual central-station practice combines both regenerative and reheat modifications to the basic Rankine cycle. For large installations, reheat makes possible an improvement of approximately 5% in thermal efficiency and substantially reduces the heat rejected to the condenser cooling water. The operating characteristics and economics of modern plants justify the installation of only one stage of reheat except for units operating at supercritical pressure.

Figure 5-3 shows a flow diagram for a 600-MW fossil-fueled reheat cycle designed for initial turbine conditions of 2520-lb/in^2 (gage) and 1000°F steam. Six feedwater heaters are supplied by exhaust steam from the high-pressure turbine and extraction steam from the intermediate and low-pressure turbines. Except for the deaerating heater (third), all heaters shown are closed heaters. Three pumps are shown: (1) the condensate pump, which pumps the condensate through oil and hydrogen gas coolers, vent condenser, air ejector, first and second heaters, and deaerating heater; (2) the condensate booster pump, which pumps the condensate through fourth and fifth heaters; and (3) the boiler feed pump, which pumps the condensate through the sixth heater to the economizer and boiler. The mass flows noted on the diagram are in pounds per hour at the prescribed conditions for full-load operation.

5.1.3 Reheat Steam Generators

The boiler designer must proportion heat-absorbing and heat-recovery surfaces to make best use of the heat released by the fuel. Waterwalls, superheaters, and reheaters are exposed to convection and radiant heat, whereas convection heat transfer predominates in air heaters and economizers.

The relative amounts of such surfaces vary with the size and operating conditions of the boiler. A small low-pressure heating plant with no heat-recovery equipment has quite a different arrangement from a large high-pressure unit operating on a reheat regenerative cycle and incorporating heat-recovery equipment.

Factors Influencing Boiler Design. In addition to the basics of unit size, steam pressure, and steam temperature, the designer must consider other factors that influence the overall design of the steam generator.

Fuels. Coal, although the most common fuel, is also the most difficult to burn. The ash in coal consists of a number of objectionable chemical elements and compounds. The high percentage of ash that can occur in coal has a serious effect on furnace performance.

At the high temperatures resulting from the burning of fuel in the furnace, fractions of ash can become partially fused and sticky. Depending on the quantity and fusion temperature, the partially fused ash may adhere to surfaces contacted by the ash-containing combustion gases, causing objectionable buildup of slag on or bridging between tubes. Chemicals in the ash may attack materials such as the alloy steel used in superheaters and reheaters.

In addition to the deposits in the high-temperature sections of the unit, the air heater (the coolest part) may be subject to corrosion and plugging of gas passages from sulfur compounds in the fuel acting in combination with moisture present in the flue gas.

Furnace. Heat generated in the combustion process appears as furnace radiation and sensible heat in the products of combustion. Water circulating through tubes that form the furnace wall lining absorbs as much as 50% of this heat, which, in turn, generates steam by the evaporation of part of the circulated water.

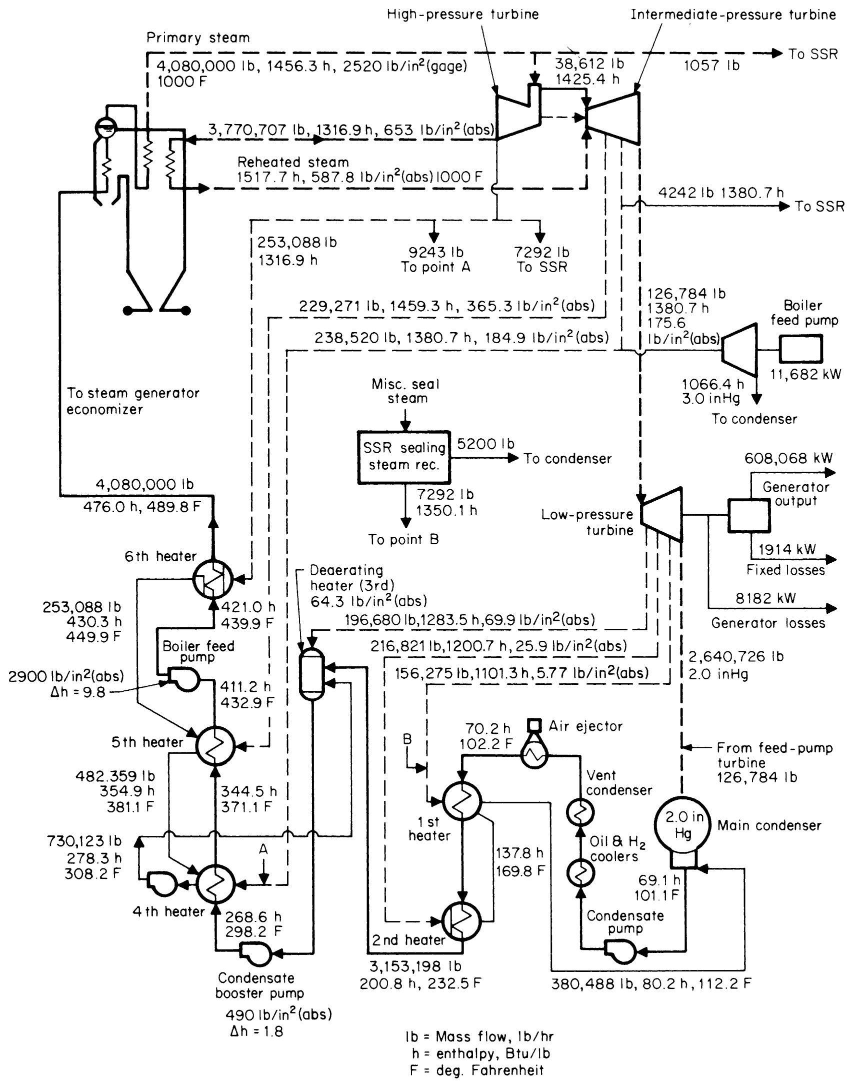


FIGURE 5-3 Reheat regenerative cycle, 600-MW subcritical-pressure fossil-fuel power plant.

Furnace design must consider water heating and steam generation in the wall tubes as well as the processes of combustion. Practically, all large modern boilers have walls comprising water-cooled tubes to form complete metal coverage of the furnace enclosure. Similarly, areas outside the furnace which form enclosures for sections of superheaters, reheaters, and economizers also use either water- or steam-cooled tube surfaces. Present practice is to use tube arrangements and configurations which permit practically complete elimination of refractories in all areas that are exposed to high-temperature gases.

Waterwalls usually consist of vertical tubes arranged in tangent or approximately so, connected at top and bottom to headers. These tubes receive their water supply from the boiler drum by means of downcomer tubes connected between the bottom of the drum and the lower headers. The steam, along with a substantial quantity of water, is discharged from the top of the waterwall tubes into the upper waterwall headers and then passes through riser tubes to the boiler drum. Here the steam is separated from the water, which together with the incoming feedwater is returned to the waterwalls through the downcomers.

Tube diameter and thickness are of concern from the standpoints of circulation and metal temperatures. Thermosyphonic (also called thermal or natural) circulation boilers generally use larger-diameter tubes than positive (pumped) circulation or once-through boilers. This practice is dictated largely by the need for more liberal flow area to provide the lower velocities necessary with the limited head available. The use of small-diameter tubes is an advantage in high-pressure boilers because the lesser tube thicknesses required result in lower outside tube-metal temperatures. Such small-diameter tubes are used in recirculation boilers in which pumps provide an adequate head for circulation and maintain the desired velocities.

Superheaters and Reheaters. The function of a superheater is to raise the boiler steam temperature above the saturated temperature level. As steam enters the superheater in an essentially dry condition, further absorption of heat sensibly increases the steam temperature.

The reheater receives superheated steam which has partly expanded through the turbine. As described earlier, the role of the reheater in the boiler is to re-superheat this steam to a desired temperature.

Superheater and reheater design depends on the specific duty to be performed. For relatively low final outlet temperatures, superheaters solely of the convection type are generally used. For higher final temperatures, surface requirements are larger and, of necessity, superheater elements are located in very high gas-temperature zones. Wide-spaced platens or panels, or wall-type superheaters or reheaters of the radiant type, can then be used. Figure 5-4 shows an arrangement of such platen and panel surfaces. A relatively small number of panels are located on horizontal centers of 5 to 8 ft to permit substantial radiant heat absorption. Platen sections, on 14- to 28-in centers, are placed downstream of the panel elements; such spacing provides high heat absorption by both radiation and convection.

Economizers. Economizers help to improve boiler efficiency by extracting heat from flue gases discharged from the final superheater section of a radiant/reheat unit (or the evaporative bank of an industrial boiler). In the economizer, heat is transferred to the feedwater, which enters at a temperature appreciably lower than that of saturated steam. Generally, economizers are arranged for downward flow of gas and upward flow of water.

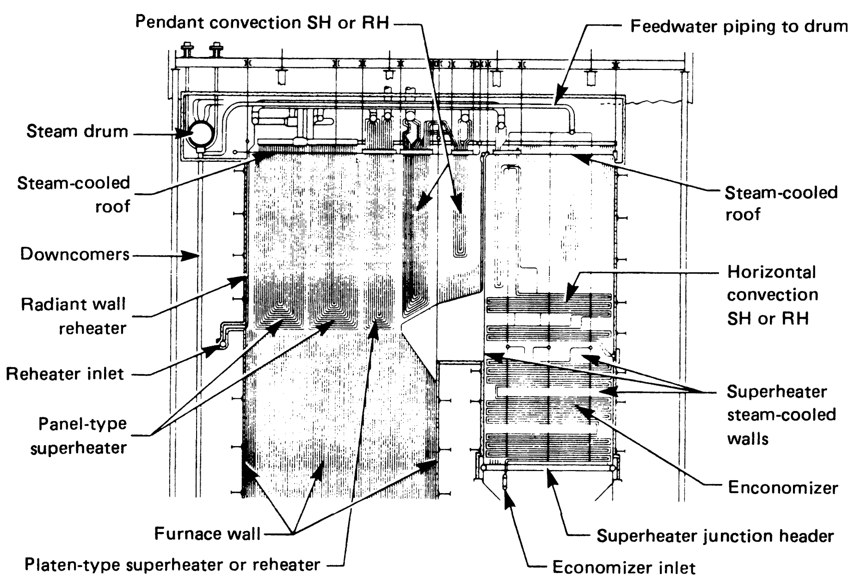


FIGURE 5-4 Arrangement of superheater, reheater, and economizer of a large coal-fired steam generator.

Water enters from a lower header and flows through horizontal tubing constituting the heating surface. Return bends at the ends of the tubing provide continuous tube elements, whose upper ends connect to an outlet header that is in turn connected to the boiler drum by means of tubes or large pipes.

As shown in Fig. 5-4, economizers of a typical utility-type boiler are located in the same pass as the primary or horizontal sections of the superheater, or superheater and reheater, depending on the arrangement of the surface. Tubing forming the heating surface is generally low-carbon steel. Because steel is subject to corrosion in the presence of even extremely low concentrations of oxygen, it is necessary to provide water that is practically 100% oxygen free. In central stations and other large plants, it is a common practice to use deaerators for oxygen removal.

Air Heaters. Steam-generator air heaters have two important and concomitant functions: they cool the gases before they pass to the atmosphere, thereby increasing fuel-firing efficiency; at the same time, they raise the temperature of the incoming air of combustion. Depending on the pressure and temperature cycle, the type of fuel, and the type of boiler involved, one of the two functions will have prime importance.

For instance, in a low-pressure gas- or oil-fired industrial or marine boiler, combustion-gas temperature can be lowered in several ways—by a boiler bank, by an economizer, or by an air heater. Here, an air heater has principally a gas-cooling function, as no preheating is required to burn the oil or gas. If the boiler is a high-pressure reheat unit burning a high-moisture subbituminous or lignitic coal, high preheated-air temperatures are needed to evaporate the moisture in the coal before ignition can take place. Here, the air-heating function becomes primary. Without exception, then, large pulverized-coal boilers either for industry or electric power generation use air heaters to reduce the temperature of the combustion products from the 600 to 800°F level to final exit-gas temperatures of 275 to 350°F. In these units, the combination air is heated from about 80°F to between 500 and 750°F, depending on coal calorific value and moisture content.

In theory, only the primary air must be heated; that is, air used to actually dry the coal in the pulverizers. Ignited fuel can burn without preheating the secondary and tertiary air. However, there is considerable advantage to the furnace heat-transfer process from heating *all* the combustion air; it increases the rate of burning and helps raise adiabatic temperature.

5.1.4 Fossil Fuels

Fossil fuels used for steam generation in utility and industrial power plants may be classified into solid, liquid, and gaseous fuels. Each fuel may be further classified as a natural, manufactured, or by-product fuel. Not mutually exclusive, these classifications necessarily overlap in some areas. Obvious examples of natural fuels are coal, crude oil, and natural gas.

Of all the fossil fuels used for steam generation in electric-utility and industrial power plants today, coal is the most important. It is widely available throughout much of the world, and the quantity and quality of coal reserves are better known than those of other fuels.

5.1.5 Classification of Coal

Coals are grouped according to rank. For the purposes of the power-plant operator, there are several suitable ranks of coal:

Anthracite
Bituminous
Subbituminous
Lignite

The following description of coals by rank gives some of their physical characteristics.

Anthracite. Hard and very brittle, anthracite is dense, shiny black, and homogeneous with no marks or layers. Unlike the lower-rank coals, it has a high percentage of fixed carbon and a low percentage of

TABLE 5-1 Classification of Coals of Rank*

	Fixed carbon limits, % (dry, mineral-matter-free basis)		Volatile matter limits, % (dry, mineral-matter-free basis)		Calorific value limits, Btu/lb (moist,† mineral-matter-free basis)		
Class and group	Equal to or greater than	Less than	Equal to or greater than	Less than	Equal to or greater than	Less than	Agglomerating character
Anthracitic							
Metaanthracite	98	...	...	2	...	...	Nonagglomerating
Anthracite	92	98	2	8	...	...	
Semianthracite‡	86	92	8	14	...	...	
Bituminous							
Low-volatile bituminous coal	78	86	14	22	...	...	
Medium-volatile bituminous coal	69	78	22	31	...	...	
High-volatile A bituminous coal	...	69	31	...	14,000§	...	Commonly glomerating¶
High-volatile B bituminous coal	...	...	...	...	13,000§	14,000	
High-volatile C bituminous coal	...	...	...	...	11,500	13,000	
					10,500	11,500	Agglomerating
Subbituminous							
Subbituminous A coal	...	...	...	...	10,500	11,500	Nonagglomerating
Subbituminous B coal	...	...	...	...	9,500	10,500	
Subbituminous C coal	...	...	...	...	8,300	9,500	
Lignitic							
Lignite A	...	...	...	...	6,300	8,300	
Lignite B	...	...	...	...	...	6,300	

Note: 1 Btu/lb = 2326 J/kg.

*This classification does not include a few coals, principally nonbanded varieties, which have unusual physical and chemical properties and which come within the limits of fixed carbon or calorific value of the high-volatile bituminous and subbituminous ranks. All of these coals either contain less than 48% dry, mineral-matter-free fixed carbon or have more than 15,500 moist, mineral-matter-free Btu per pound.

†Moist refers to coal containing its natural inherent moisture but not including visible water on the surface of the coal.

‡If agglomerating, classify in low-volatile group of the bituminous class.

§Coals having 69% or more fixed carbon on the dry, mineral-matter-free basis shall be classified by fixed carbon, regardless of calorific value.

¶It is recognized that there may be nonagglomerating varieties in these groups of the bituminous class, and there are notable exceptions in the high-volatile C bituminous group.

Source: ASTM Standards D388, *Classification of Coals by Rank.*

volatile matter. Anthracites include a variety of slow-burning fuels merging into graphite at one end and into bituminous coal at the other. They are the hardest coals on the market, consisting almost entirely of fixed carbon, with the little volatile matter present in them chiefly as methane, CH_4. Anthracite is usually graded into small sizes before being burned on stokers. The "metaanthracites" burn so slowly as to require mixing with other coals, while the "semianthracites," which have more volatile matter, are burned with relative ease if properly fired. Most anthracites have a lower heating value than the highest-grade bituminous coals. Anthracite is used principally for heating homes and in gas production.

Some semianthracites are dense, but softer than anthracite, shiny gray, and somewhat granular in structure. The grains have a tendency to break off in handling the lump, and produce a coarse, sand-like slack. Other semianthracites are dark gray and distinctly granular. The grains break off easily in handling and produce a coarse slack. The granular structure has been produced by small vertical cracks in horizontal layers of comparatively pure coal separated by very thin partings. The cracks are the result of heavy downward pressure, and probably shrinkage of the pure coal because of a drop in temperature.

Bituminous. By far the largest group, bituminous coals derive their name from the fact that on being heated they are often reduced to a cohesive, binding, sticky mass. Their carbon content is less than that of anthracites, but they have more volatile matter. The character of their volatile matter is more complex than that of anthracites, and they are higher in calorific value. They burn easily, especially in pulverized form, and their high volatile content makes them good for producing gas. Their binding nature enables them to be used in the manufacture of coke, while the nitrogen in them is utilized in processing ammonia.

The low-volatile bituminous coals are grayish-black and distinctly granular in structure. The grain breaks off very easily, and handling reduces the coal to slack. Any lumps that remain are held together by thin partings. Because the grains consist of comparatively pure coal, the slack is usually lower in ash content than are the lumps.

Medium-volatile bituminous coals are the transition from high-volatile to low-volatile coal and, as such, have the characteristics of both. Many have a granular structure, are soft, and crumble easily. Some are homogeneous with very faint indications of grains or layers. Others are of more distinct laminar structure, are hard, and stand handling well.

High-volatile A bituminous coals are mostly homogeneous with no indication of grains, but some show distinct layers. They are hard and stand handling with little breakage. The moisture, ash, and sulfur contents are low, and the heating value is high.

High-volatile B bituminous coals are of distinct laminar structure; the layers of black, shiny coal alternate with dull, charcoal-like layers. They are hard and stand handling well. Breakage occurs generally at right angles and parallel to the layers, so that the lumps generally have a cubical shape.

High-volatile C bituminous coals are of distinct laminar structure, are hard, and stand handling well. They generally have high moisture, ash, and sulfur contents and are considered to be free-burning coals.

Subbituminous. These coals are brownish black or black. Most are homogeneous with smooth surfaces, and with no indication of layers. They have high moisture content, as much as 15% to 30%, although appearing dry. When exposed to air they lose part of the moisture and crack with an audible noise. On long exposure to air, they disintegrate. They are free-burning, entirely noncoking, coals.

Lignite. Lignites are brown and of a laminar structure in which the remnants of woody fibers may be quite apparent. The word *lignite* comes from the Latin word *lignum* meaning wood. Their origin is mostly from plants rich in resin, so they are high in volatile matter. Freshly mined lignite is tough, although not hard, and it requires a heavy blow with a hammer to break the large lumps. But on exposure to air, it loses moisture rapidly and disintegrates. Even when it appears quite dry, the moisture content may be as high as 30%. Owing to the high moisture and low heating value, it is not economical to transport it long distances.

Unconsolidated lignite (B in Table 5-1) is also known as "brown coal." Brown coals are generally found close to the surface, contain more than 45% moisture, and are readily won by strip mining.

5.1.6 Impact of Fuel on Boiler Design

The most important item to consider when designing a utility or large industrial steam generator is the fuel the unit will burn. The furnace size, the equipment to prepare and burn the fuel, the amount of heating surface and its placement, the type and size of heat-recovery equipment, and the flue-gas-treatment devices are all fuel dependent.

The major differences among those boilers that burn coal or oil or natural gas result from the ash in the products of combustion. Firing oil in a furnace results in relatively small amounts of ash; there is no ash from natural gas. For the same output, because of the ash, coal-burning boilers must have larger furnaces and the velocities of the combustion gases in the convection passes must be lower. In addition, coal-burning boilers need ash-handling and particulate-cleanup equipment that costs a great deal and requires considerable space.

TABLE 5-2 Representative Coal Analyses

	Medium-volume bituminous	High-volume bituminous	Subbituminous C	Low-sodium lignite	Medium-sodium lignite	High-sodium lignite
Total H_2O, %	5.0	15.4	30.0	31.0	30.0	39.6
Ash, %	10.3	15.0	5.8	10.4	28.4	6.3
VM, %	31.6	33.1	32.6	31.7	23.2	27.5
FC, %	53.1	36.5	36.6	26.9	18.4	26.6
Btu/lb, as fired	13,240	10,500	8,125	7,590	5,000	6,523
Btu/lb, MAF	15,640	15,100	12,650	12,940	12,020	12,050
Fusion (reducing), °F						
Initial def.	2,170	1,990	2,200	2,075	2,120	2,027
Softening	2,250	2,120	2,250	2,200	2,380	2,089
Fluid	2,440	2,290	2,290	2,310	2,700	2,203
Ash analysis, %						
SiO_2	40.0	46.4	29.5	46.1	62.9	23.1
Al_2O_3	24.0	16.2	16.0	15.2	17.5	11.3
Fe_2O_3	16.8	20.0	4.1	3.7	2.8	8.5
CaO	5.8	7.1	26.5	16.6	4.8	23.8
MgO	2.0	0.8	4.2	3.2	0.7	5.9
Na_2O	0.8	0.7	1.4	0.4	3.1	7.4
K_2O	2.4	1.5	0.5	0.6	2.0	0.7
TiO_2	1.3	1.0	1.3	1.2	0.8	0.5
P_2O_5	0.1	0.1	1.1	0.1	0.1	0.2
SO_3	5.3	6.0	14.8	12.7	4.6	17.7
Sulfur, %	1.8	3.2	0.3	0.6	1.7	0.8
Lb H_2O/million Btu	3.8	14.7	36.9	40.8	60.0	60.7
Lb ash/million Btu	7.8	14.3	7.1	13.7	56.8	9.7
Fuel-fired,* 1000 lb/h	405	520	705	750	1,175	900

Note: 1 Btu/lb = 2326 J/kg; $t_{°C} = (t_{°F} - 32)/1.8$; 1 lb = 0.4536 kg; 1 Btu = 1055 J.
*Constant heat output, nominal 600-MW unit, adjusted for efficiency.

Table 5-2 lists the variation in calorific values and moisture contents of several coals, and the mass of fuel that must be handled and fired to generate the same electrical-power output. These values are important because the quantity of fuel required helps determine the size of the coal-storage yard, as well as the handling, crushing, and pulverizing equipment for the various coals.

Furnace Sizing. The most important step in coal-fired unit design is to properly size the furnace. Furnace size has a first-order influence on the size of the structural-steel framing, the boiler building and its foundations, as well as on the sootblowers, platforms, stairways, steam piping, and duct work. The fuel-ash properties that are particularly important when designing and establishing the size of coal-fired furnaces include

The ash fusibility temperatures (both in terms of their absolute values and the spread or difference between initial deformation temperature and fluid temperature)

The ratio of basic to acidic ash constituents

The iron/calcium ratio

The fuel-ash content in terms of pounds of ash per million British thermal units

The ash friability

These characteristics and others translate into the furnace sizes in Fig. 5-5, which are based on the six coal ranks shown in Table 5-2. This size comparison illustrates the philosophy of increasing the furnace plan area, volume, and the fuel burnout zone (the distance from the top fuel nozzle to the furnace arch), as lower-grade coals with poorer ash characteristics are fired.

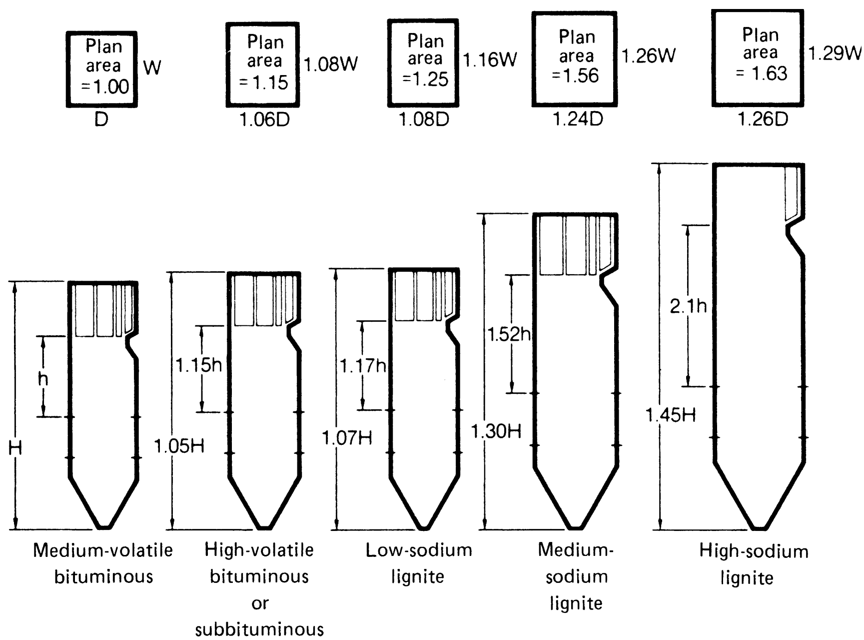


FIGURE 5-5 Effect of coal rank on furnace sizing (constant heat output).

Figure 5-5 is a simplified characterization of actual furnaces built to burn the fuels listed in Table 5-2. Wide variations exist in fuel properties within coal ranks, as well as within several subclassifications (e.g., subbituminous A, B, C), each of which may require a different size furnace.

Among the most important design criteria in large pulverized-fuel furnaces are net heat input in British thermal units per hour per square foot of furnace plan area (NHI/PA) and the vertical distance from the top fuel nozzle to the furnace arch. Furnace dimensions must be adequate to establish the necessary furnace retention time to properly burn the fuel as well as to cool the gaseous combustion products. This is to ensure that the gas temperature at the entrance to the closely spaced convection surface is well below the ash-softening temperature of the lowest-quality coal burned. Heat-absorption characteristics of the walls are maintained using properly placed wall blowers to control the furnace outlet gas temperature by removing ash deposited on the furnace walls below the furnace outlet plane.

5.1.7 Environmental Considerations

Concerns for the control of air quality have probably had the largest single impact on power plant site selection, design, operation, and cost. The three classes of emissions which are of major concern are nitrogen oxides, sulfur oxides, and particulate matter.

Nitrogen Oxides. In the United States, nitrogen oxides can be controlled within federal, state, and local regulatory limits by in-furnace and postcombustion techniques. With respect to firing systems, each steam-generator manufacturer has developed specific design concepts for reducing nitrogen oxides. The common characteristics of all of these designs, however, included a careful regulation of the fuel/air ratio in the firing zone where the major fraction of the fuel nitrogen compounds are liberated and control of the heat-liberation pattern in the furnace. Postcombustion reduction methods utilizing reagents with or without catalysts are somewhat similar in concept among the steam-generator suppliers.

Particulate Control. The traditional particulate control device in power plant applications has been the electrostatic precipitator. In recent years, fabric filters (also called "baghouses") have become increasingly popular.

In electrostatic precipitation, suspended particles in the gas are electrically charged, then driven to collecting electrodes by an electrical field; the electrodes are rapped to cause the particles to drop into collecting hoppers. This process differs from mechanical or filtering processes in which forces are exerted directly on the particulates rather than the gas as a whole. Effective separation of particles can be achieved with lower power expenditure, with negligible draft loss, and with little or no effect on the composition of the gas.

The principle of electrostatic precipitation is relatively simple. The process applies an electrostatic charge to dust particles with a corona discharge and passes them through an electric field where the particles are attracted to a collecting surface. The basic elements of a precipitator include a source of unidirectional voltage, corona or discharge electrodes, collecting electrodes, and a means of removing the collected matter.

Single-stage (Cottrell-type) precipitators combine the ionizing and collecting step. In the more common plate type, the electrodes are suspended between plates on insulators connected to a high-voltage source. A voltage differential created between the discharge and collecting electrodes develops a strong electric field between them. The flue gas is passed through the field and a unipolar discharge of gas ions, from the discharge electrode, is attached to the particulate matter.

5.1.8 Fabric Filtration

Fabric filters, or baghouses, have a long history of applications in both dry and wet filtration processes to recover chemicals or control stack emissions. Available materials limited early baghouse installations to temperatures below 250°F, and air dilution was frequently used ahead of the baghouse. In addition, the chemical-resistance characteristics of the bags also curtailed fabric filtration. These two limitations retarded its development for many years, particularly as available precipitator equipment met the existing regulations.

Serious consideration of this technology began after 1970; interest heightened as installations on large coal-fired boilers demonstrated good operating characteristics and high particulate-removal efficiencies.

5.1.9 Flue-Gas Desulfurization Systems

Flue-gas desulfurization (FGD) began in England in 1935. The technology remained dormant until the mid-1960s when it became active primarily in the United States and Japan. Since then, over 50 FGD processes have been developed, differing in the chemical reagents and the resultant end products.

The most common FGD system is a lime/limestone wet scrubber. After the flue gas has been treated in the precipitation (or baghouse), it passes through the induced fans and enters the SO_2 scrubber. If the required SO_2 removal efficiency is less than 85%, a fraction of the flue gas can be treated while bypassing the rest to mix with and reheat the saturated flue gas leaving the scrubber.

For higher-sulfur fuels requiring SO_2 removal efficiencies of 90% or greater, the entire flue-gas stream must be treated. Upon leaving the SO_2 absorption section, the flue gas is passed through entrainment separators to remove any slurry droplets mixed with the gas. The saturated flue gas is then reheated approximately 25 to 50°F above the water dewpoint before it is vented to the stack.

For low- to medium-sulfur fuels, an alternate scrubbing technology is dry scrubbing. This process minimizes water consumption and eliminates the requirement for flue-gas reheating but requires more expensive additives than the wet limestone systems.

The typical dry SO_2 absorber is a cocurrent classifying spray dryer. Flue gas enters the top of the absorber through inlet assemblies containing swirl vanes. The absorbent is injected pneumatically into the center of each swirler assembly by ultrasonic atomizing nozzles that require an air pressure of about 60 lb/in^2 (gage). Slurry feed pressures are 10 to 15 lb/in^2 (gage). The compressed air induces primary dispersion of the absorbent slurry by mechanical shear forces produced by the two fluid streams. Final dispersion is accomplished by shattering the droplets with ultrasonic energy produced by the compressed air used with a proprietary nozzle design. Then ultrasonic nozzles generate extremely fine droplets, which have diameters that range from 10 to 50 μm, as shown by photographic studies.

The flue-gas outlet design requires that effluent gases make a 180° turn before leaving the absorber. Besides eliminating product accumulation in the outlet duct, the abrupt directional change also allows the larger particles to drop out in the absorber product hopper. This design curtails the particulate loading to the fabric filter. Consequently, the number of cleaning cycles as well as abrasion of the filter medium are reduced.

As compared with ordinary fly-ash collection applications, fabric filters together with dry scrubbing offer a broader choice of design options. In conventional fly-ash collection applications, the fabric filter experiences flue-gas temperatures about 100 to 150°F higher than encountered in dry scrubbing. Filter media unsuitable at the higher temperatures can be used when the fabric filter follows a dry absorber. In particular, acrylic fibers become attractive because of their strength and flex characteristics, as well as their ability to support more vigorous cleaning methods like mechanical shaking.

5.1.10 Advanced Methods of Using Coal

Coal, which is the most abundant and economically stable fossil fuel in the United States, continues to grow in use while under pressure to meet the most stringent federal and local emissions requirements. This trend has added to the cost and complexity of coal combustion technologies.

Emission-control methods that facilitate the use of coal in power plants can be classified as

Precombustion processes

In situ combustion processes

Postcombustion processes

Precombustion processes include methods to clean the coal of sulfur-bearing compounds by wet separation, coal gasification, and coal liquefaction techniques. Coal gasification involves the partial oxidation of coal to produce a clean gas or by production of a "clean fuel" through coal liquefaction. Sulfur and ash are removed in these processes. The use of coal to produce a gas is not a new idea; it has been used to produce "town gas" for over 200 years. But its use in the United States had almost disappeared by 1930, because natural gas was abundant and low in cost. Concerns about the availability and economic stability of gas supplies, along with environmental trends, have renewed interest in coal gasification to produce substitute natural gas (SNG) and low- and medium-heat-content (LBTU and MBTU) gas for chemical feedstock or power plant fuel. Coal gasification in the combined-cycle mode has been well established as a viable technology for producing power with very low emissions both in the United States and Europe. New plants are using technologies such as high-temperature gas turbines, hot-gas cleanup to remove 99% of the sulfur (H_2S), and higher-pressure combined steam cycles to achieve overall efficiencies of greater than 40%. New integrated gasification combined-cycle (IGCC) plants of as much as 250 MWe are available. IGCC technology produces very low emissions per kilowatt of power and is therefore very attractive for the production of power. Likewise, coal liquefaction is not a new technology, but is only in limited commercial use in the United States. South Africa is the largest producer of synthetic liquid fuels from coal. Large-scale production of synthetic liquid fuels from coal began in 1910 in Germany with the Fischer-Tropsch process, which is used to produce a variety of fuels.

In fluidized-bed combustion, an in situ combustion-emission-control process, 90% to 95% of the SO_2 is captured during combustion by a sorbent (limestone). In this process, the NO_x production is low because of the low temperature at which the combustion reaction takes place. NO_x levels well fired below 0.25 lb/MBtu have been achieved with certain coals. Fluidized-bed combustion was developed in the 1950s and is now available for electric power plants of up to 300-MWe size. The technology has three distinct types of units: bubbling bed, hybrid velocity, and circulating fluidized bed (CFB). CFB technology is the most popular fluidized-bed process and has evolved as a low-emission technology with excellent fuel flexibility for the production of power. Bubbling and hybrid-velocity fluidized-bed technologies have demonstrated low emissions while burning low-rank coals, waste fuels such as petroleum coke, and renewable fuel such as wood and peat. Hybrid-velocity fluidized-bed combustion can be readily retrofit to many older boilers that need pollution-control technology. Pressurized fluidized-bed combustion is used to achieve low sulfur and NO_x emissions of fluidized-bed combustion integrated with a gas turbine to achieve high cycle efficiency, and therefore make more efficient use of coal.

Postcombustion control processes are widely used for the capture of sulfur and particulate. Lime-based scrubbers for SO_2 removal and equipment for particulate control were described in Sec. 5.1.9.

Processes and equipment for removal of NO_x from flue gases leaving boilers have been widely used in Europe and are being applied in the United States. In situ control of NO_x by modifications to firing technology and over-fire air can reduce NO_x as much as 50%. Selective noncatalytic control (SNCR) involves ammonia or urea sprayed in the proper place in the boiler to reduce NO_x. More NO_x reduction can be achieved by selective catalytic reduction (SCR), which uses ammonia in a postcombustion control system. SCR can reduce NO_x levels well below those from a conventional pulverized-coal boiler.

Coal gasification is an efficient way to produce electric power while minimizing the emissions from the combustion of coal. Coal gasification can achieve cycle efficiencies above 40% when the gas turbine cycle is completely integrated with the steam cycle. This is referred to as the *integrated gasification combined cycle* (IGCC) (Fig. 5-6). In an IGCC plant, the gas from the gasification process is burned in a boiler or gas turbine for the generation of electric power. The process also uses the heat from the gas turbine exhaust to produce electric power from a steam cycle.

In the gasification process, coal is partially reacted with a deficiency of air to produce low-heating-value fuel gas. The gas is cleaned of particulate and then sulfur compounds in a hot-gas cleanup system. Elemental sulfur is disposed of or sold.

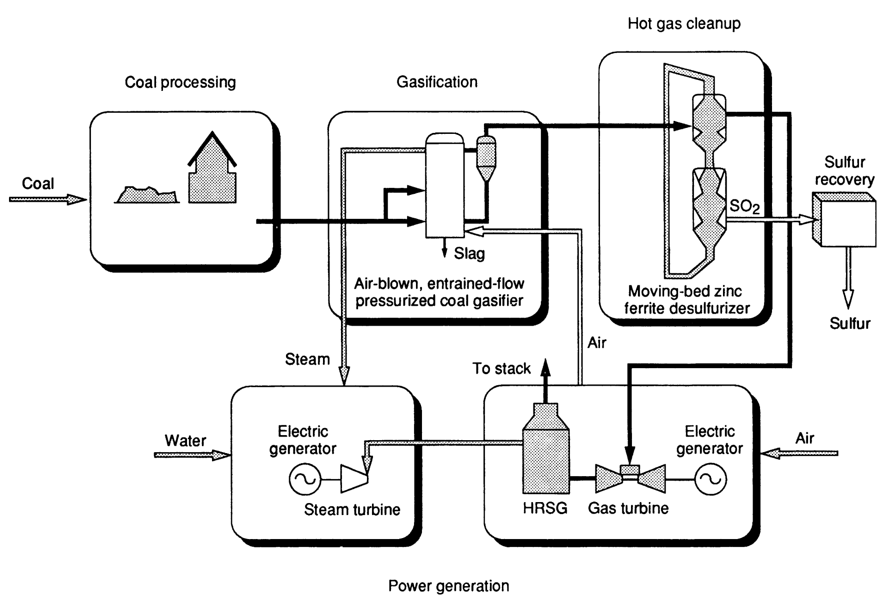


FIGURE 5-6 Integrated combined-cycle power plant.

5.1.11 Fluidized-Bed Combustion

For decades, fluidized-bed reactors have been used in noncombustion reactions in which the thorough mixing and intimate contact of the reactants in a fluidized bed result in high product yield with improved economy of time and energy. Although conventional methods of burning coal can also generate energy with very high efficiency, fluidized-bed combustion can burn coal efficiently at a temperature low enough to avoid many of the problems of conventional combustion.

The outstanding advantage of fluidized-bed combustion (FBC) is its ability to burn high-sulfur coal in an environmentally acceptable manner without the use of flue-gas scrubbers. A secondary benefit is the formation of lower levels of nitrogen oxides compared to other combustion methods.

5.1.12 Circulating Fluidized-Bed Steam Generators

Figure 5-7 shows a typical CFB steam generator. Crushed fuel and sorbent are fed mechanically or pneumatically to the lower portion of the combustor. Primary air is supplied to the bottom of the combustor through an air distributor, with secondary air fed through one or more elevations of air ports in the lower combustor. Combustion takes place throughout the combustor, which is filled with bed material. Flue gas and entrained solids leave the combustor and enter one or more cyclones where the solids are separated and fall to a seal pot. From the seal pot, the solids are recycled to the combustor. Optionally, some solids may be diverted through a plug valve to an external fluidized-bed heat exchanger (FBHE) and back to the combustor. In the FBHE, tube bundles absorb heat from the fluidized solids.

Bed temperature in the combustor is essentially uniform and is maintained at an optimum level for sulfur capture and combustion efficiency by heat absorption in the walls of the combustor and in the FBHE (if used). Flue gas leaving the cyclones passes to a convection pass, air heater, baghouse, and induced-draft (ID) fan. Solids inventory in the combustor is controlled by draining hot solids through an ash cooler.

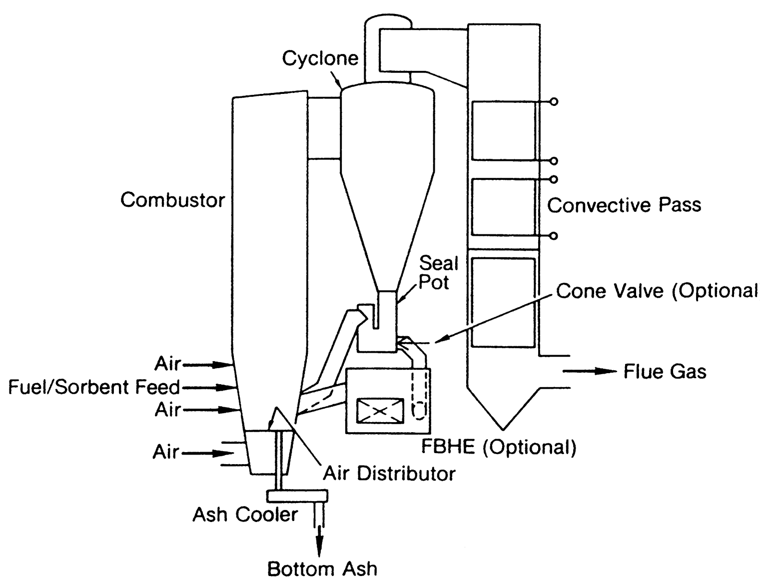


FIGURE 5-7 Typical circulating fluidized-bed (CFB) steam generator.

5.2 NUCLEAR POWER PLANTS

By GEORGE H. MILEY

5.2.1 Nuclear Energy

Introduction. The United states is the world's largest supplier of commercial nuclear power. In 2005, there were 104 U.S. commercial nuclear generating units that were fully licensed to operate. One reactor, however, Brown's Ferry unit 1 has been shut down since 1985. Therefore, some sources cite only 103 units. Together, they provide about 20% of the nation's electricity—second only to coal as a fuel source.

The Energy Information Administration (EIA) reports that the U.S. nuclear industry generated 788,556 million kilowatt hours of electricity in 2004 (Fig. 5-8), a new U.S. (and international) record. Although no new U.S. nuclear power plants have come on line since 1996, this is the industry's fifth annual record since 1998.

General. Applying the nuclear process for electrical production involves consideration of characteristics substantially different from those associated with the use of fossil fuels. With fossil fuels or with hydro, the amount of energy source (fuel) supplied to the power plant is proportional to the power demanded at that time. With nuclear power, however, the fuel for a substantial amount of energy output is physically located in the converter at any time. A second important characteristic of the nuclear process is the energy density. The thermal energy density in a typical fossil boiler (heated volume or core volume) is in the range of 0.20 kW/L; in a typical nuclear power generator it is in

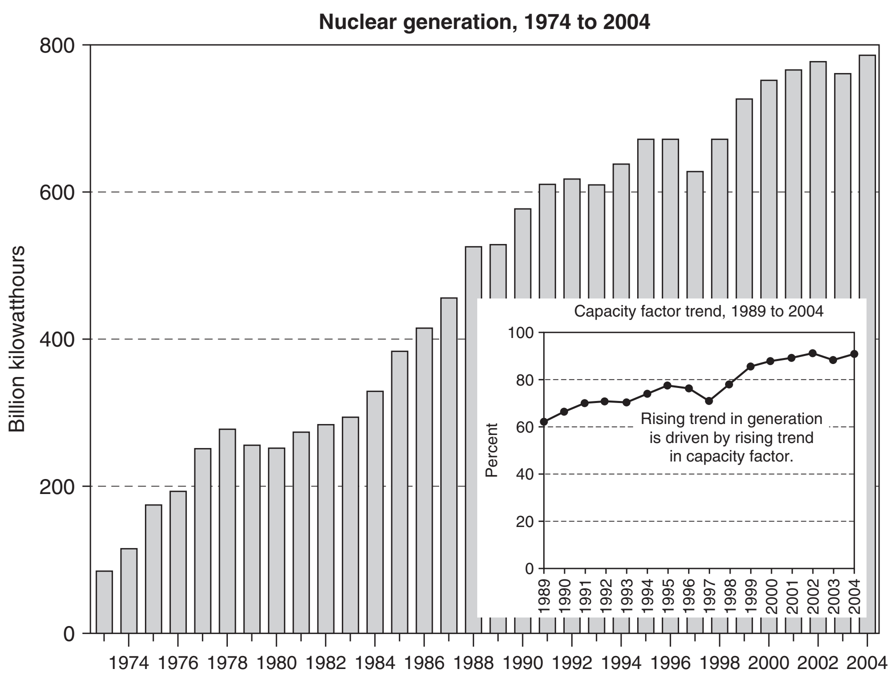


FIGURE 5-8 U.S. nuclear power generation. (***Source:*** Energy Information Administration, *Monthly Energy Review.*)

the range of 80 kW/L. A third important difference is that of continued low-level heat generation (decay heat) when the nuclear process is shut down following power operation. A fourth important difference is that of emanations. The fossil process requires the intake of large volumes of air and fuel and the corresponding exhaust of large volumes of waste gas, including CO_2, SO_3, NO_2, etc., some particulate matter, and in the case of coal-fired boilers, substantial quantities of ash. The nuclear process, however, requires only the input of the material placed in the core; its output is fuel-element materials plus radioactive "waste" products from the fission process. This residue includes small quantities of gases which may be released or may be stored and solids which are contained within the fuel.

These and other more subtle aspects introduce many new considerations in the equipping and regulation of the nuclear process.

5.2.2 Mass-Energy Relationships

One of the first applications of the special theory of relativity proposed by Einstein in 1905 was the interrelation between mass and energy, expressed by the equation $E = mc^2$. Thus, a change in nuclear mass appears as energy. If the mass m is expressed in kilograms and the velocity of light c in meters per second, the energy E is in joules.

$$\begin{aligned} E(\text{J}) &= \text{mass(kg)} \times (2.998 \times 10^8 \text{ m/s})^2 \\ &= \text{mass(kg)} \times 8.99 \times 10^{16} \text{ m}^2\text{s}^2 \end{aligned} \tag{5-1}$$

The amounts of energy involved in single nuclear events are usually very small. Thus, for convenience, the electronvolt (the energy acquired by any charged particle carrying a unit electronic charge falling through a potential of 1 V) is often used. One electronvolt (eV) $= 1.602 \times 10^{-19}$ J and, correspondingly, 1 keV $= 1.602 \times 10^{-16}$ J. One MeV $= 1.602 \times 10^{-13}$.

The mass-energy relationships become

$$\begin{aligned} E(\text{eV}) &= \text{mass(kg)} \times \frac{8.99 \times 10^{16} \text{ m}^2/\text{s}^2}{1.602 \times 10^{-19} \text{ J/eV}} \\ &= \text{mass(kg)} \times 5.61 \times 10^{35} \text{ eV/kg} \\ E(\text{keV}) &= \text{mass(kg)} \times 5.61 \times 10^{32} \text{ keV/kg} \\ E(\text{MeV}) &= \text{mass(kg)} \times 5.61 \times 10^{29} \text{ MeV/kg} \end{aligned}$$

where $1 \text{ J} = 1 \text{ m}^2 \cdot \text{kg/s}^2$.

It is often convenient to use the energy corresponding to 1 atomic mass unit (amu). One amu $= 1.657 \times 10^{-27}$ kg (1 amu $= {}^1/_{12}$ of the mass of a neutral atom of ^{12}C).

$$\begin{aligned} E_{\text{amu}} &= 1.66 \times 10^{-27} \text{ kg} \times 5.61 \times 10^{29} \text{ MeV/kg} \\ &= 931 \text{ MeV/amu} \end{aligned} \tag{5-2}$$

The atomic mass of a nuclide can be evaluated in terms of the masses of its constituent particles and the *binding energy* (Fig. 5-9). The mass of the nuclide is less than the sum of its constituent particles in the free state. If ΔM is the decrease in mass when a number of protons, neutrons, and electrons combine to form an atom, then the mass-energy equivalence principle states that an amount of energy equal to $\Delta E = c^2 \Delta M$ is released in the process. The difference in mass ΔM is called the *mass defect*; it is the amount of mass which would be converted to energy if a particular atom or nuclide were to be assembled from the requisite number of protons, neutrons, and electrons. The same amount of energy would be needed to break the atom into its constituent particles, and the energy equivalent of the mass defect is therefore a measure of the binding energy of the nuclei. The mass of

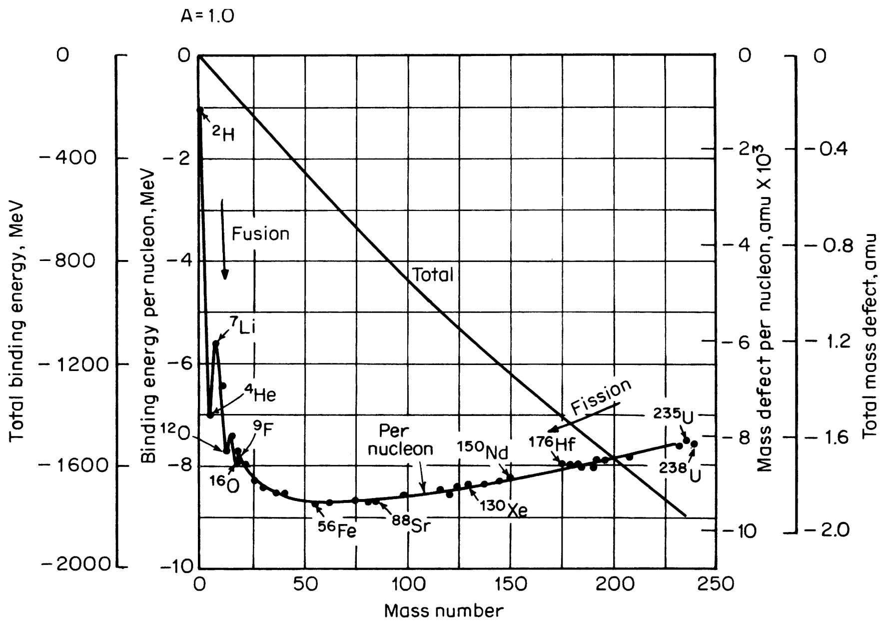


FIGURE 5-9 Mass defects and binding energies of nuclei.

the constituent particles is the sum of Z proton masses, Z electrons, and $A - Z$ neutrons, where A refers to the mass number of the element. Pairs of protons and electrons can be represented by hydrogen atoms; the loss in mass which accompanies the formation of the hydrogen atom from the proton and an electron is negligible. The mass defect can then be written $\Delta M = ZM_H + (A - Z)M_n - M_{ZA}$, where M_H is the mass of the hydrogen atom, 1.008142 amu; M_n is the mass of the neutron, 1.008982 amu; and M_{ZA} is the mass of nuclide of concern.

Figure 5-9 provides an approximate picture of the nuclear binding energy. In the higher mass numbers, the actual binding energy is not the same for each particle in the nucleus. After the maximum of the curve, almost every successive particle (proton or neutron) is bound less tightly than those already present, and the overall average decreases. The binding energy represented, however, is sufficiently accurate for engineering evaluations.

5.2.3 The Fission Process

In the higher mass numbers, several of the naturally occurring elements are radioactive or have a characteristic which enables them to emit nuclear particles and be transmuted to different elements as a function of time. The various naturally occurring series are designated the thorium, uranium, and actinium series. These designations are related to the elements at or near the head of the series and can be expressed as multiples of a number N, where N is an integer. The series are indicated by $4N$, $4N + 2$, and $4N + 3$, respectively. There is no naturally occurring $4N + 1$ element; such an element has been created in the process of artificial nuclear transmutation. This element is designated neptunium and has the mass characteristic of $4N + 1$. It, too, heads a radioactive series. The four radioactive series are shown in Fig. 5-10.

A number of elements with high mass numbers, both natural and artificially produced, undergo a process of nuclear fission. In the fission process, a nucleus absorbs a neutron and the resulting compound nucleus is so unstable that it immediately breaks up into parts. As shown by the arrow labeled "fission" in Fig. 5-9, the fission products have a lower mass and larger binding energy, resulting in

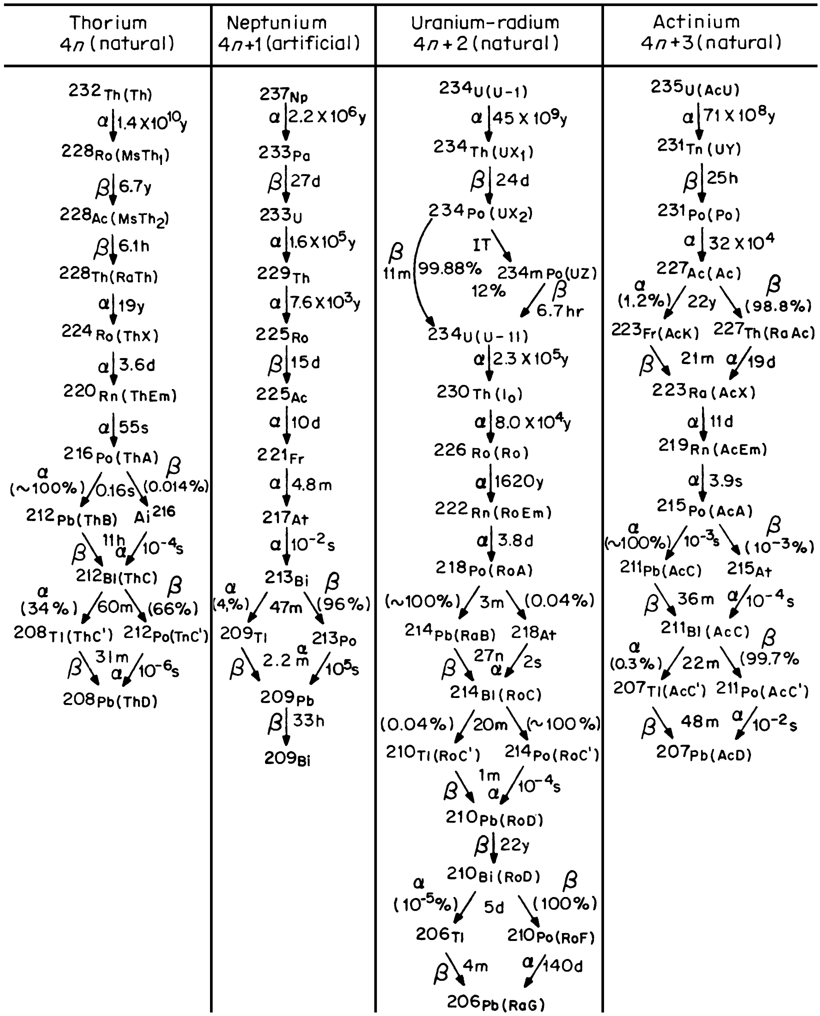


FIGURE 5-10 The four radioactive series.

a release of energy in the form of kinetic energy of the products. (Also note in Fig. 5-9 that the arrow for "fusion" shows lighter elements fusing together to create higher mass products, again with a release of energy by emission of high-speed products. Many of the heavy nuclides can be induced to fission, but most only with neutrons of high energy. Naturally occurring heavy nuclides that fission with neutrons of energy in the range of the neutrons produced by the fission are uranium isotopes ^{235}U and ^{238}U and thorium 232. In addition, artificially produced nuclides ^{233}U and ^{239}Pu, produced by (n, β) reactions in ^{232}Th and ^{238}U, respectively, are capable of fission. The fission process, in a nuclear reactor, is initiated by neutrons which are generated as part of the process. The general fission process may be expressed by

$$^{m}F + {}^{1}n \rightarrow {}^{x}A + {}^{m-(x+C)}B + C\,{}^{1}n \qquad (5\text{-}3)$$

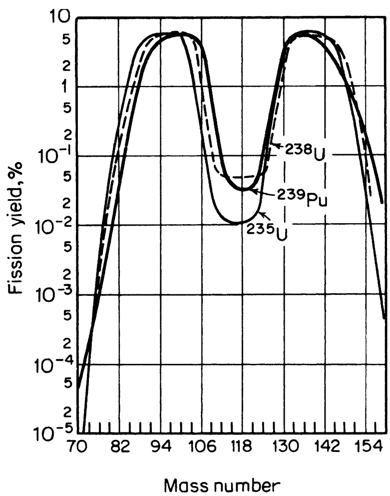


FIGURE 5-11 Fission yield.

where F = fuel nuclide, mass number m
n = neutron
A, B = fragment nuclides
C = number of neutrons produced
x = atomic number

The percentage of nuclide production as a function of mass number is shown in Fig. 5-11.

A typical example is the fission of ^{235}U with the production of two most likely fission fragments.

$$^{235}\text{U} + {}^{1}n \rightarrow {}^{95}A + {}^{139}B + 2\,{}^{1}n \qquad (5\text{-}4)$$

The mass balances of this equation are

Before fission		After fission	
^{235}U	– 235.124 amu	^{95}A	–94.945 amu
^{1}n	1.009 amu	^{139}B	138.955 amu
		$2\,{}^{1}n$	2.018 amu
236.133 amu			235.918 amu

TABLE 5-3 Distribution of Fission Energy

Energy	MeV
Kinetic energy of fission fragments	168 ± 5
Instantaneous gamma-ray energy	5 ± 1
Kinetic energy of fission neutrons	5 ± 0.5
Beta particles from fission products	7 ± 1
Gamma rays from fission products	6 ± 1
Neutrinos	~10
Total fission energy	201 ± 6

The mass change resulting from fission is 236.133 − 235.918 = 0.215 amu, which by the relationship of mass to energy is equivalent to E(J) = mass(amu) × 1.49 × 10^{-10} J/amu, which represents ~3.2 × 10^{-11} J/fission or approximately 200 MeV/fission (or 3.2 × 10^{-11} W · s/fission).

The major portion of this energy is released immediately as kinetic energy of the fission fragments, the fission neutrons, and instantaneous gamma rays. A portion of the energy is released gradually from the decay of the fission fragments. Table 5-3 shows the distribution of fission energy. For practical purposes, the neutrino energy, because of the low probability of interaction of neutrinos with matter, is not recoverable. (This leaves about 190 MeV, or 3.0 × 10^{-11} J, recoverable per fission.)

5.2.4 Neutron Interaction

Each neutron interacting with a nucleus does not always result in fission; some are scattered and some are involved in radiative capture, that is, initiate the radiation of other particles and/or photons to reduce the target atom to a stable state. The neutron-absorption probability for a given nuclide is referred to as its cross section and is expressed in units of area. Since very small areas are involved, the special unit for cross section is the barn, equal to 10^{-24} cm^2.

The cross section for a neutron interaction varies with energy. An explanation is that, quantum mechanically, the wavelength λ of the neutron is inversely proportional to its energy E or velocity, and may be expressed by

$$\lambda = \frac{2.86 \times 10^{-8}}{\sqrt{E(\text{ev})}} \quad \text{mm} \qquad (5\text{-}5)$$

For fast neutrons (about 1 MeV), λ is of the order of 10^{211} mm, and for thermal neutrons (about 0.03 MeV), λ is about 1.7 3 10^{27} mm. The slower neutrons behave as though they had a diameter approaching that of the atom, and thus have a larger probability of interaction.

5.2.5 Radiation

Nuclide Composition. The elements of the periodic table, both naturally occurring and artificial, are composed of protons (except for hydrogen), neutrons, and electrons. Many of the elements have two or more isotopic forms, states which have the same atomic number but a different atomic mass because of a different number of neutrons in the nucleus.

Most of the naturally occurring elements are stable, that is, do not eject particles to change to a different isotope or a different element. However, some naturally occurring elements, as indicated in Fig. 5-10, are conditionally stable and have a probability for transmutation. Out of the total number of atoms present, the probability indicates that a certain number of the atoms will, by ejecting a particle, change to an isotope or a new element. The mode of decay for a given isotope is predictable. The pattern is sometimes complex and follows a decay chain.

Radioactive Transmutation. For every radioactive material, there are characteristic quantities that may be used to describe the process. Each radioactive nuclide has a definite probability of decaying in unit time. This decay probability has a constant value, characteristic of the particular radioisotope. In a given sample, the rate of decay at any instant is proportional to the number of radioactive atoms present at that time. If N is the number of radioactive atoms present at time t and λ is the decay constant, the decay rate is given by $dn/dt = -\lambda N$ for a simple decay scheme. Integrating this over the interval N_0 to N gives

$$N = N_0 e^{-\lambda t} \tag{5-6}$$

where N = number of atoms remaining unchanged at any time t
N_0 = initial number of atoms
λ = disintegration constant

The reciprocal of the decay constant 1/λ is the mean or average life of the radioactive species = t_m. A more widely used quantity for quantifying radioactive decay is the half-life, that period of time during which half the atoms originally present are transmuted. If N is set equal to $\frac{1}{2}N_0$ and the above equation is solved for t, the value becomes

$$t_{1/2} = \frac{\ln 2}{\lambda} = \frac{0.6931}{\lambda} \tag{5-7}$$

In a radioactive species, a nuclide may undergo successive decay before reaching the ground state. For a compound decay scheme involving two states A and B, the net rate of change of B with time is given by

$$\frac{dN_b}{dt} = \lambda_A N_A - \lambda_B N_B \tag{5-8}$$

where the solution is $N = [\lambda_A N_{A0}/(\lambda_B - \lambda_A)]\,(e^{-\lambda_A t} - e^{-\lambda_B t})$. The first term on the right represents the production of B from the decay of A; the second term is the decay of B. N_{A0} is the number of parent atoms at time $t = 0$. Sample decay curves in Fig. 5-12 show both a simple decay and a compound two-stage decay.

If the radiation occurs by the emission of a quantity of energy (photon), the nuclide retains its atomic weight and number. If the decay occurs by emission of a particle, the nuclide changes to an isotope (same atomic number), an isobar (same mass number), or a different element.

Artificial elements, including those resulting from the fission process, are very likely to be radioactive. In some cases, this activity results in the emission of a photon of energy to allow the atom to reach a lower energy state. In other cases, a particle is emitted; the particle emitted for some decaying nuclides is a neutron. These delayed neutrons are important to the regulation of the fission process.

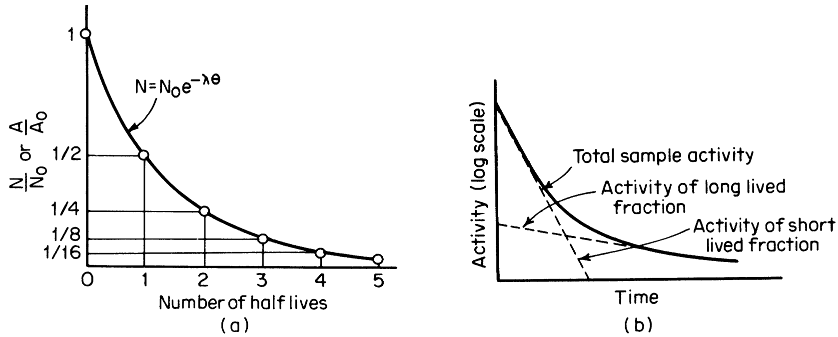


FIGURE 5-12 (*a*) Radioactive decay of a single radionuclide as a function of half-life; (*b*) decay of mixture of independent radionuclides.

Types of Radiation. There are three categories of radiation emanations of biological concern in nuclear power. The first category is that of charged particles, principally alpha particles and beta particles. The second is that of uncharged particles, chiefly neutrons. The third is that of photons or gamma rays. The charged particles directly produce ionization by collision with neutral atoms. Neutrons and photons indirectly produce ionization by liberating directly ionizing particles or by initiating nuclear transformations.

In radioactivity, a conventional unit is the curie, that quantity of any radioactive material giving 3.7×10^{10} disintegrations/s. For small quantities of radiation, the millicurie and the microcurie, 3.7×10^{7} and 3.7×10^{4} disintegrations/s, respectively, are frequently used. The rutherford (rd), equal to 10^{6} disintegrations/s, is sometimes used. The SI unit of radioactivity is the becquerel [1 curie (Ci) = 3.7×10^{10} becquerel (Bq)].

The radioactivity, the decay constant, and the weight are related by

$$\frac{dN}{dt} = \lambda N = \frac{\lambda W_A}{G_w} \tag{5-9}$$

where λ = decay constant, disintegrations/s
W = weight of the material, g
A = Avogadro's number = 6.02×10^{23} atoms/mol
G_w = gram atomic weight of the material, g/mol
N = number of atoms

This equation shows that a given amount of radioactivity may occur from a large mass with a small decay rate or a small mass which has a high decay rate.

Radiation dosage is expressed in four ways:

1. *Absorbed dose* (D), which is the energy absorbed per unit mass at a specific place in a material. The standard of absorbed dose is the gray; 1 Gy = 1 J/kg. The special unit of absorbed dose is the rad = 0.01 J/kg = 0.01 Gy. A subset is the absorbed-dose index, which is the maximum absorbed dose, at a point, within a 300-mm-diameter sphere centered at the point and consisting of material equivalent to soft tissue with a density of 1 g/cm^3.
2. *Dose equivalent* (H). In general, the biological equivalent of a given absorbed dose depends on the type of radiation and the irradiation conditions. The product of modifying factors, assigned to weigh the effect on a given organ, and the absorbed dose is the dose equivalent. The special unit of H is the rem (where D is in rads, H is in rems). A subset of this is the dose-equivalent index, which is the maximum dose equivalent, at a point, within a 300-mm-diameter sphere centered at the point and consisting of material equivalent to soft tissue with a density of 1 g/cm^3.

3. *Kerma* (K), which is the sum of the initial kinetic energies of the charged particles produced by indirectly ionizing radiation per unit mass of the material in which the interaction takes place. The units of K are grays or rads.
4. *Exposure* (X) is the measure of a particular field of electromagnetic radiation (x- or gamma rays) to ionize air. The special unit of exposure is the roentgen (R) $= 2.58 \times 10^{-4}$ coulombs (C)/kg of air.

5.2.6 Nuclear Plant Safety

The nuclear-powered steam supply system characteristics of substantial energy potential present in the reactor, radiation production during the fission process, and continued radiation production and heat generation after shutdown require that special safety precautions be taken in design and operation of a nuclear plant. The health and welfare of the public depends on both the continuation of the plant's power production and the avoidance of any incident which would endanger the environment. In order to achieve the latter goal and to aid the former, special regulations relating to nuclear plants have been formulated.

Workers in power plants are covered by federal regulations, with special attention being devoted to radiation protection. A guiding principle applied throughout the industry is known as ALARA. This directs management and workers to seek *as low an exposure as is reasonably achievable*. Application of ALARA requires careful planning and a balance between minimizing exposure versus work requirements.

5.2.7 Federal Regulations

Title 10 of the Code of Federal Regulations (10 CFR) has the following parts which are of primary importance to nuclear power facilities:

Part 20, Standards for Protection against Radiation

Part 50, Licensing of Production and Utilization Facilities

Part 55, Operators' Licenses

Part 70, Special Nuclear Material

Part 100, Reactor Site Criteria

There are several other parts of 10 CFR which relate to the usage or handling of radioactive material. Most of the parts previously listed have appendixes which treat requirements for specific subjects. Authority for regulation of commercial, nuclear-powered plants is vested with the U.S. Nuclear Regulatory Commission. This authority includes the licensing of new facilities and the surveillance of operating facilities. An applicant for a nuclear-powered plant is required to apply for a license to construct and operate the facility. Such application includes the submission of Safety Analysis Reports which describe the design bases, the design, and the analyses performed to show that plant performance and conditions will be within established limits.

5.2.8 Standards

Appendix A of 10 CFR Part 50 provides general design criteria for nuclear power plants. Criterion 1 requires that structures, systems, and components important to safety be designed, fabricated, erected, and tested to quality standards. The nuclear standards program of the American National Standards Institute has developed a sizable group of standards for this requirement. The principal design, systems, and operation standards are those developed by ASME, IEEE, ANS, and ISA.

Many other documents providing criteria, standard practices, or guidance are available. In the nuclear area, specific designs have not been repeated frequently enough to accumulate a significant backlog of experience. As a result, many of the "standards" are developed to provide leadership in addressing given areas.

The Nuclear Regulatory Commission provides guidance in many areas of design, construction, and operation through Regulatory Guides. Individual guides may cite a standard as an acceptable method of addressing the area concerned.

5.2.9 Quality Assurance

The best defense against incidents which endanger the public is to prevent them. In a similar way, the best system performance is effected when malfunctions are eliminated. Reliability is the interface between quality assurance and safety. Reliability can neither be tested nor legislated into equipment; it must be built in. High quality in design, procurement, installation, and operation will lead to a system that has high availability, good reliability, and a low probability of incurring an accident. Quality assurance is a total systems approach to achieving these aims. Quality assurance does represent an increase in costs; this increase must be balanced against safer operation and savings resulting from less time lost, fewer repairs, and better control. The prime responsibility for an effective quality assurance program lies with the owner/operator of the plant, who may delegate portions of the program to major suppliers.

5.2.10 Nuclear Energy System

Reactor-System Assembly. To achieve a self-sustaining, but regulated fission process, and for the energy released to be extracted and converted to electricity, a reactor system is constructed.

The nuclear fuel, usually uranium, is fabricated into fuel elements. The typical design for the fuel of a light-water power reactor involves the fuel in oxide form. Where the fuel is uranium, the uranium dioxide is fabricated into pellets, right circular cylinders approximately 19 mm high and 8 mm in diameter.

In light-water reactors, the uranium dioxide material is typically enriched to a low value, approximately 3% to 7%, in the fissionable isotope ^{235}U. This enrichment is necessary because light water has an appreciable neutron-absorption cross section. The extra neutrons available from the added fissile material compensate for the absorption in the moderator. The fuel-pellet material is usually of a ceramic nature (e.g., UO_2); the pellets are dished at both ends to allow for differential thermal expansion and fuel volumetric growth with burnup.

The pellets are inserted into fuel tubes, typically thin-walled tubes of stainless steel or Zircalloy. An open space (with the column of pellets spring-loaded) at the top of the tube is provided to accommodate generation of gases during the fission process. The tubes are sealed top and bottom and are assembled into a configuration involving fixed spacing in a fuel assembly. A representative fuel assembly is shown in Fig. 5-13. This assembly has an overall length of approximately 4.5 m with an active length of approximately 3.8 m.

Plutonium, ^{239}Pu, is produced in the fuel elements during power operation by the absorption of neutrons in the ^{238}U. This material is fissionable and may be recovered during fuel reprocessing and fabricated into new fuel elements. However, to date, regulations in the United States have prevented fuel reprocessing, largely due to concerns about proliferation of materials for use in weapons. Consequently, spent fuel from U.S. plants is currently being stored, awaiting future decisions about reprocessing. In contrast, reprocessing plants have been constructed in Europe.

In gas-cooled reactors in the United States, the fuel-element design differs from that of light-water reactors. The recent gas-cooled reactor elements are hexagonal graphite blocks into which blind longitudinal holes are drilled to receive rods of fuel particles. The fissile material is enriched uranium carbide, UC_2. Kernels of this material are coated with a pyrolytic carbon–silicon carbide–pyrolytic carbon sandwich. Fertile material in the form of thorium oxide, ThO_2, kernels is also used. A fertile material is one which, by absorption of neutrons, is changed to a material which can be fissioned. In this case, ^{232}Th is converted to ^{233}U, which has superior characteristics for fission reactions. The kernels are coated with two layers of pyrolitic carbon. The two types of fuel particles are mixed in the proper proportions and are formed with a carbon matrix into fuel “rods” about 15.6 mm in diameter and about 60 mm long. These rods are inserted into the holes in the graphite blocks. Through holes are provided in the block for the helium coolant flow. These loaded graphite blocks or “fuel assemblies” form the basic module for the core of the gas reactor.

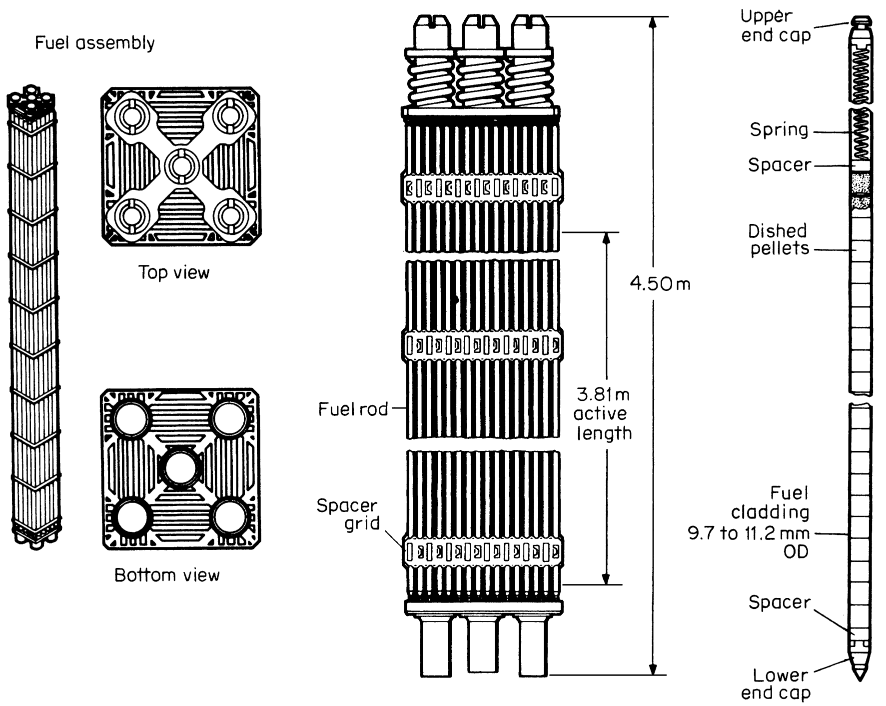


FIGURE 5-13 PWR fuel assembly.

The required number of fuel assemblies to produce a power output desired for the reactor plant are assembled into a reactor-core configuration approximating a right circular cylinder. This configuration provides a high volume-to-surface ratio which minimizes the neutron leakage and conserves the neutrons produced for further fission action. For a 1300-MW (electrical) light-water nuclear plant, a representative core assembly might involve 241 fuel assemblies each weighing approximately 660 kg, for a core equivalent diameter of 3.6 m, a core height of approximately 5 m, and a total core weight of approximately 160 metric tons.

Control-Element Assemblies. In each fuel assembly, several holes are shown. These open holes are spaces into which control elements are inserted for regulation of the fission process. The individual control elements may be grouped typically into control-element assemblies. Control elements for current Pressurized Water Reactors (PWRs) (Fig. 5-22) are located in the fuel elements in this fashion. Control elements of Boiling Water Reactors (BWRs) are blade-type cruciform units. These units are inserted into or withdrawn from the spaces between the fuel assemblies. The control-element assemblies are selected from a material which absorbs neutrons; therefore, by insertion into the fuel assembly or withdrawal from the fuel assembly, the amount of neutrons available for fission production can be reduced or increased, respectively, as required for reactor-system performance. These control-element assemblies are inserted or withdrawn by electromechanical or hydraulic drive mechanisms.

Moderator and Heat-Transfer Medium. The thermal energy released from the core must be conveyed to the electric generator at a rate and in a fashion which meets the requirements. Some consideration has been given to the use of a reactor core to heat gas which is supplied to a magnetohydrodynamic generator. However, commercial systems for the present and near future will continue to use steam turbines as the motive power for the generator.

If fuel of low enrichment (e.g., 3% to 4%) in ^{235}U or other fissile isotope is used, a moderator is needed to take advantage of the larger cross section at thermal neutron energy. If enrichments greater than 20% are used, sufficient fissile material is present to overcome the nonfission capture effects of ^{238}U, and a moderator is not needed. In this case, the reactor is said to be "fast" (referring to the neutron velocity), and liquid metals (selected for low neutron absorption) are used as a coolant.

5.2.11 Plant Arrangement

Primary-System Configuration. The fuel elements, assembled into the core arrangement, are positioned within a reactor vessel by support structures also referred to as reactor internals. The reactor vessel also contains, guides, and directs the primary coolant. Elementary configurations are shown in Fig. 5-14.

In the BWRs, the reactor vessel also contains the steam-separation apparatus, since the coolant is converted to steam in the core. Steam is piped from the reactor vessel to the turbine and condensate

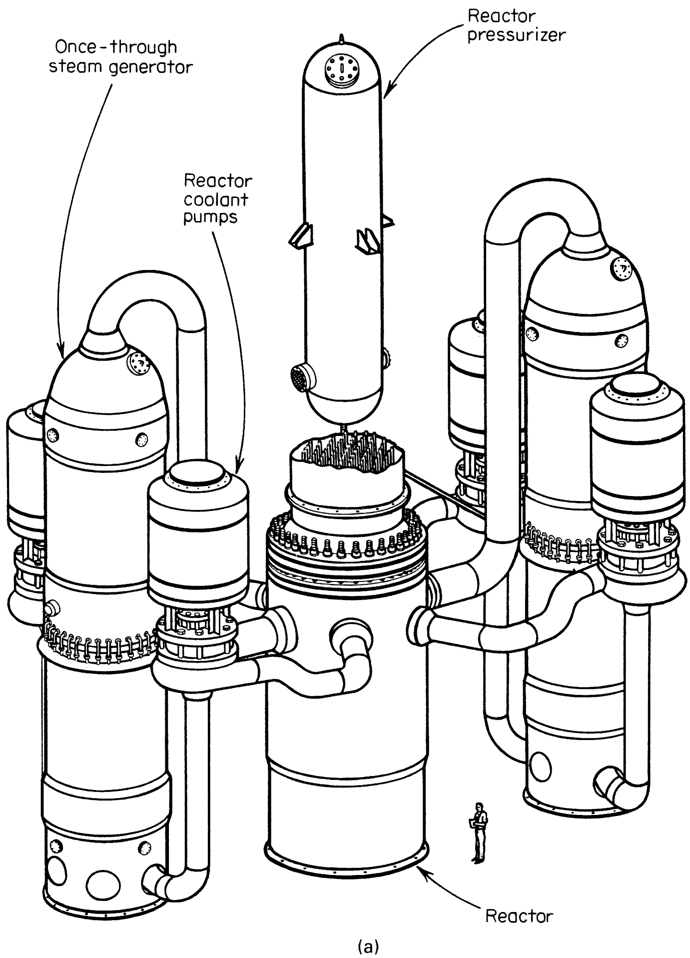
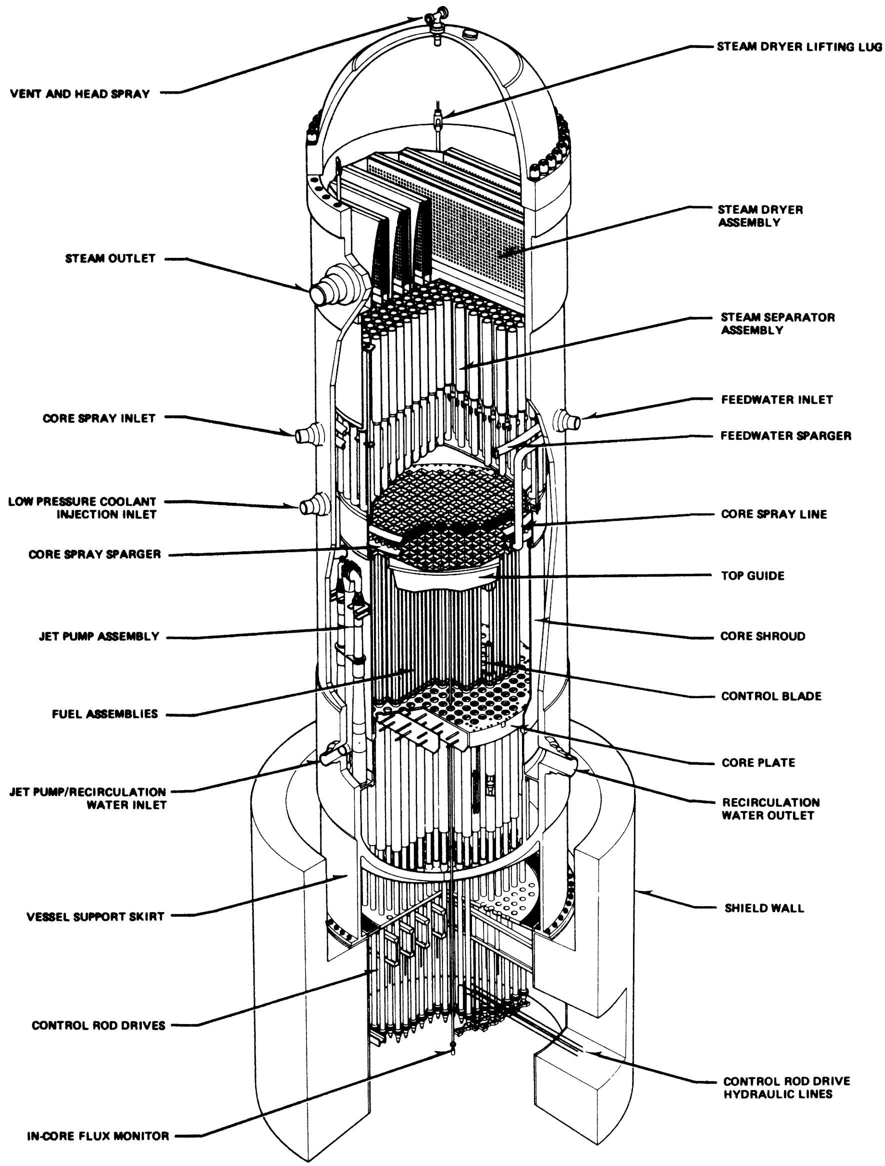


FIGURE 5-14 (*a*) PWR arrangement with once-through steam generators; (*b*) BWR arrangement. (*General Electric Co.*)



(b)

FIGURE 5-14 (*Continued*)

is returned from the hotwell through feedwater systems to the reactor vessel. Recirculation loops on the reactor vessel increase the recirculation ratio to effect better steam separation and provide better control of void fraction within the core (reactivity control). Access to the core for control rods and for in-core instrumentation is provided from the bottom of the reactor vessel to avoid the steam-separation area at the top.

In PWRs, the primary system is maintained at a subcooled condition by operating at a pressure greater than saturation. Conventionally this pressure is in the range of 12 to 16 MPa. This pressure is maintained by a pressurizer connected to the primary piping. A steam-water interface is maintained in the pressurizer by the action of electric resistance heaters which boil water to raise the pressure or spray flow to quench steam and lower the pressure. The reactor vessel is connected, by heavy piping, to one or more steam generators, and coolant is circulated through this primary system by large pumps. On the secondary side of the steam generator are located the customary steam piping complex, the turbine condenser, and feedwater system.

In liquid-metal reactors, the concerns associated with coolant metal–water reaction in the event of leakage and the induced activity in the primary metal sometimes direct the design to an intermediate heat-exchange system between the primary-metal coolant and the steam system.

Containment Structure. To provide a secondary barrier for radioactivity in the fluid systems and a fourth barrier for the activity in the fuel (the fuel matrix or ceramic, the fuel sheath, and the primary-system boundary are the other three), the primary-system components are located within a containment structure. This structure is designed to confine gases and materials that might occur in the event of an inadvertent release from the primary barrier(s).

The principal types of containments that have been used are

1. The steel sphere made of sections of steel sheet welded together
2. The "inverted lightbulb" of the BWR, which is a concrete cavity in the shape described with a pressure-retaining steel liner
3. The domed cylinder for the PWR; the cylinder of reinforced concrete with an impervious internal steel membrane
4. The cylindrical prestressed-concrete reactor vessel (PCRV) for the GCRs

These structures serve to provide isolation and shielding for the primary-system components. Instrumentation, control, and electric power conductors must pass through the pressure seal while maintaining its integrity. To provide this capability, banks of containment penetrations are provided. The requirements for these items are described by IEEE 317, *Electric Penetration Assemblies in Containment Structures for Nuclear Power Generating Stations.*

5.2.12 Plant Operations

Nuclear Plant Costs. A large amount of equipment and capital investment is required for a nuclear plant. A nuclear steam-supply system (NSSS) is arranged and equipped to initiate, sustain, and regulate the fission process in the reactor fuel, transfer heat from the fuel to the steam generators, and produce steam to supply a turbine. Auxiliary fluid systems to provide chemical cleaning and conditioning of the primary-system fluid, control of chemical shim (if used), supply and purification of secondary water (if required), and processing of radioactive effluents (liquid and gas) are associated with the NSSS. Equipment for instrumentation, control, protection, and electric power distribution is included. Typical equipment supplied in a nuclear steam-supply package is shown in Table 5-4.

This large investment in equipment encourages the development of nuclear plants of large capacity to reduce the per megawatt cost. Plants in the current generation have been from 500 to 1300 MW in electrical capacity. Depending on the size of the system, a single nuclear plant can represent 10% to 20% or more of the system operating capacity. This large size, the high capital investment, plus a low fuel cost of the nuclear plant direct the base loading of the nuclear plant.

Performance Evaluation. Evaluating the nuclear steam-supply system for performance as a source of steam energy often requires that the system be modeled, that is, the system equations be developed. The principal components whose characteristics are important for transient analysis are the reactor fission process and thermal process and the primary coolant piping and the steam generator(s) in a PWR along with the associated control and protection systems.

The functions of these components, in combination, are (1) the reactor core through the fission process converts potential energy in the uranium atoms to thermal energy in the fuel elements; (2) in a

TABLE 5-4 Typical Equipment in a Nuclear Steam-Supply Package

Reactor system:	Rod-drive servicing equipment
Pressure vessel and internals	Vessel-servicing equipment
Control rods	Auxiliary systems, PWR:
Rod-drive mechanisms	Chemical and volume control
Recirculation pumps, BWR only	Residual-heat removal
Steam generators } PWR only	Spent-fuel pit
Pressurizer } PWR only	Safety-injection system
Controls and instrumentation:	Component-cooling system
Reactor controls and instrumentation	Radioactive-waste processing
Rod-drive position indicators	Sampling system
In-core instrumentation	Auxiliary systems, BWR:
Ex-core nuclear instrumentation	Reactor-water cleanup
Plant protection system	Standby-liquid control
Plant monitoring and supervisory system	Core-isolation cooling
Steam bypass control equipment	Residual-heat removal
Auxiliary system controls and instrumentation	Core spray system
Control-room panels	High-pressure coolant injection
Feedwater regulating equipment	Fuel-pool cooling and filtering
Tools and servicing equipment	Radioactive-waste processing
Fuel-handling equipment	

PWR, GCR, LMFBR, the primary coolant piping conveys the thermal energy to the steam generator(s) and the steam generator(s) transfer the energy from the primary fluid to the secondary fluid. Since the secondary fluid is maintained at a lower pressure, boiling is introduced and steam is generated; (3) in a BWR, the directly produced steam is separated; (4) the PWR pressurizer maintains pressure on the primary coolant to assure that the primary thermal process takes place in a subcooled condition; and (5) the BWR recirculation system maintains a circulation of primary fluid sufficient to ensure adequate steam separation and to regulate the void fraction in the core (reactivity control).

Neutron Multiplication. The description of performance of the reactor begins with the neutron kinetics, the time behavior of the fission neutrons. A generation of neutrons begins with a neutron flux density from the previous generation (or from the source). A fuel nuclide absorbs a neutron and probably undergoes fission. A small percentage of absorbed neutrons do not cause fission; so the number of neutrons released per capture of a neutron is given by

$$\eta = \nu \frac{\Sigma_f}{\Sigma_u} \tag{5-10}$$

where η = number of neutrons released per capture
ν = number of neutrons released per fission
Σ_f = cross section of the fuel for fission
Σ_u = absorption cross section (fission and nonfission) in the fuel

In a reactor assembly where the fissions are initiated principally by thermal neutrons, some fissioning will be introduced by the fast neutrons before they have been thermalized. The ratio of the total number of fast neutrons produced by neutrons of all energies to the number produced by thermal neutrons is given by ϵ.

During the slowing-down (thermalizing) process, some neutrons are captured in nonfission processes; the fraction escaping such capture is ρ. When the neutrons have been thermalized, they will diffuse in the core region until absorbed in the fuel or in some other material, structure, moderator, poison, etc. The fraction absorbed in the fuel is given by

$$f = \frac{\text{thermal neutrons absorbed in fuel}}{\text{total thermal neutrons absorbed}}$$

If n neutrons are present in one generation (in an infinite system), the multiplication factor, k_{inf}, or the ratio of the neutrons in one generation to those in the next generation, is given by

$$k_{inf} = \frac{n\eta\,\epsilon p f}{n} = \eta\,\epsilon p f$$

This is also known as the four-factor formula.

In a reactor system of finite dimensions, there is also leakage out of the system, so that the infinite multiplication factor must be adjusted to provide for leakage. Considering the diffusion process and the boundary conditions, it is evident that for a reactor of finite physical dimensions, there will be some leakage (loss) of neutrons from the boundary. In describing the process of slowing down the neutrons, two quantities are developed. The first of these is the diffusion length L, which is equal to one-sixth of the net vector distance that a monoenergetic neutron travels from its source to the point where it is absorbed by a nucleus. A second quantity is the buckling B, which represents the "bending" or appreciable reduction of the value of neutron flux at any point in the reactor.

These quantities may be used to develop two factors which take into account the finite size of the reactor and the leakage which occurs. For the first factor, the term $e^{-B^2L_s}$ represents the nonleakage probability of the neutrons as they slow down, where L_s is a slowing-down length. The algebraic loss, by diffusion, of thermal neutrons in a volume element is $-D\nabla^2\phi = DB^2\phi$. The ratio of thermal leakage to thermal absorption is

$$\frac{\text{Thermal leakage}}{\text{Thermal absorption}} = \frac{DB^2\phi}{\Sigma_\alpha\phi} = L^2B^2 \quad \text{so } \frac{D}{\Sigma_\alpha} = L^2 \tag{5-11}$$

Adding the thermal absorption to the thermal leakage, effectively adding unity to both sides of the equation, and inverting, gives, for the second factor, the ratio

$$\frac{\text{Thermal absorption}}{\text{Thermal leakage} + \text{thermal absorption}} = \frac{1}{1 + L^2B^2}$$

which accounts for the nonleakage probability at thermal energy.

For a finite reactor then, the effective multiplication factor may be expressed by a combination of these two factors:

$$k_{eff} = k_{inf}\frac{e^{-B^2L_s}}{1 + L^2B^2} \tag{5-12}$$

When fission occurs, more than 99% of the resulting neutrons are produced within 10^{-3} s. The remaining neutrons are produced during the decay of the fission fragments. The time required for their production varies; they may be separated into groups for convenience. These delayed neutrons are essential to the regulation of the fission process.

Reactor Kinetics. For development of control equations, a single delayed group model (Fig. 5-15) may be used for approximation of the neutron production.

The production of neutrons for fission initiation including generation and thermalization is given by $K_{inf}\Sigma_\alpha\phi e^{-B^2L_s}$, where ϕ is the neutron population, Σ_α is the absorption cross section, and $e^{-B^2L_s}$ is the nonleakage probability during thermalization.

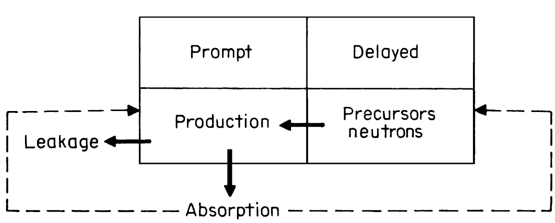


FIGURE 5-15 Single-neutron delayed group model.

The leakage of neutrons is $D\nabla^2\phi = -DB^2\phi$. Assuming β is the fraction of neutrons delayed, the neutron balance for the main group is

$$\dot{n} = (1 - \beta)k_{\text{inf}}\Sigma_\alpha \phi e^{-B^2L_s} + \lambda C - DB^2\phi - \Sigma_\alpha \phi \tag{5-13}$$

where λ = decay constant of the delayed neutron precursors
C = population of delayed neutron precursors
ϕ = neutron fluence rate (flux)

Since $\phi = n\nu$, the equation becomes

$$\dot{n} = n\left[(1 - \beta)k_{\text{inf}}\,\nu\Sigma_\alpha e^{-B^2L_s} - \nu DB^2 - \nu\Sigma_\alpha\right] + \lambda C$$

But

$$k_{\text{eff}} = \frac{k_{\text{inf}}\,e^{-B^2L_s}}{1 + L^2B^2}\frac{1}{\bar{l}} = \nu\Sigma_\alpha(1 + L^2B^2)$$

and $D = \Sigma_\alpha L^2$, where $\bar{l}$ is the average lifetime of the neutrons.

Substituting for k_{inf} and D gives

$$\dot{n} = n[(1 - \beta)k_{\text{eff}}\nu\Sigma_\alpha(1 + L^2B^2) - \nu\Sigma_\alpha(1 + L^2B^2)] + \lambda C$$

Substituting $1/\bar{l}$ for $\nu\Sigma_\alpha(1 + L^2B^2)$ gives

$$\dot{n} = n\left[(1 - \beta)\frac{k_{\text{eff}}}{\bar{l}}\frac{1}{\bar{l}}\right] + \lambda C$$

$$= n\frac{(1 - \beta)k_{\text{eff}} - 1}{\bar{l}} + \lambda C$$

The balance equation for the delayed group is

$$\dot{c} = n(\beta k_{\text{inf}}\,\nu\Sigma_\alpha e^{-B^2L_s}) - \lambda c \tag{5-14}$$

Substituting

$$\frac{k_{\text{inf}}\,e^{-B^2L_s}}{1 + L^2B^2} = k_{\text{eff}}$$

gives

$$\dot{c} = n[\beta k_{\text{eff}}\,\nu\Sigma_\alpha(1 + L^2B^2)] - \lambda c$$

and substituting $1/\bar{l} = \nu\Sigma_\alpha(1 + L^2B^2)$ gives

$$\dot{c} = n\frac{\beta k_{\text{eff}}}{\bar{l}} - \lambda c$$

Rearranging the equation for $\dot{n}$ gives

$$\dot{n} = n\frac{(k_{\text{eff}} - 1) - \beta k_{\text{eff}}}{\bar{l}} - \lambda\dot{C} = n\frac{k_{\text{eff}}(1 - \beta) - 1}{\bar{l}} - \lambda C$$

Since k_{eff} is very close to 1, $\beta k_{\text{eff}} \approx \lambda\beta$ and reactivity ρ is the ratio of the excess multiplication factor to the effective multiplication or

$$\rho = \frac{k_{\text{eff}} - 1}{k_{\text{eff}}} = \frac{k_{\text{ex}}}{k_{\text{eff}}} \quad \text{or} \quad \frac{\delta k}{k_{\text{eff}}}$$

Where the deviations from criticality are small, $k_{eff} \approx 1$ and

$$\rho \approx k_{ex} = k_{eff} - 1 = \delta k \qquad \text{so } \dot{n} = n[(\delta k - \beta)/\bar{l}] - \lambda C$$

The power level P is proportional to the neutron concentration. There are typically six delayed neutron groups. The balance equations then become

$$\dot{P} = \frac{P(\delta k - \beta)}{\bar{l}} + \sum_{j=1}^{\phi} \lambda_j C_j \tag{5-15}$$

$$\dot{C}_j = P\beta_j/\bar{l} - \lambda_{jC_j} \tag{5-16}$$

where j represents the delayed neutron group and β_j, λ_j, C_j are the fraction, decay constant, and concentration of delayed neutrons, respectively, of the jth group.

The balance equations show that, in the steady state, the effective multiplication factor is equal to 1. If the K_{eff} increases above 1, the multiplication will increase with time; if K_{eff} decreases below 1, the multiplication will decrease below 1. If $k_{eff}(1 - \beta) - 1$ increases to a value of 1 or greater, the reactor is said to be prompt critical and the rate of power increase depends on the ratio of $k_{eff}(1 - \beta) - 1$ to l. Since l is so small (about 10^{-4} s or less for a thermal reactor; 10^{-7} s or less for a fast reactor), P increases very rapidly with time for any appreciable value of $k_{eff}(1 - \beta) - 1$. Regulation of the process at these rates with conventional apparatus is very difficult. For this reason, k_{eff} in power reactors is kept below the value $1/(1 - \beta)$ when the reactor is operating.

Reactivity Control. The reactivity is affected by neutron absorption. The absorption occurs principally from control-element-type absorbers, dissolved-chemical control absorbers, resonance absorption in the fuel, absorption in the moderator, and absorption by fission products. The absorption initiated by the control elements and the chemical shim are varied by the operators and are characterized by δk_c.

The reactivity effect from the fuel is caused by the widening of the resonance peaks (Fig. 5-16) with temperature, which increases the nonfission capture of neutrons. With cores containing large amounts of ^{238}U and ^{232}Th, this Doppler effect is negative; that is, increasing the power level introduces a reactivity change which opposes the increase. The Doppler coefficient varies with coolant and fuel temperature and with moderator voids. Typical values of the Doppler coefficient are shown in Fig. 5-16. The curves may be approximated by the equation

$$\delta k_F = AT_F - BT_F^2 \tag{5-17}$$

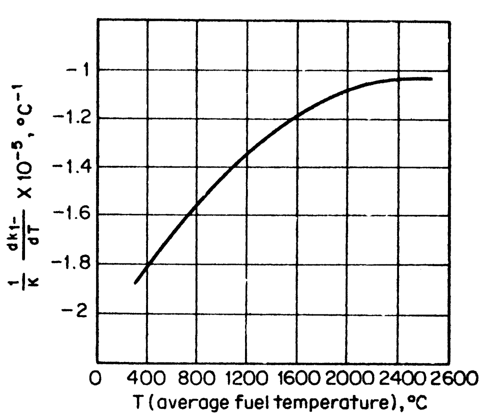


FIGURE 5-16 Values of the Doppler coefficient of reactivity. (*General Electric Co.*)

The reactivity effect of the moderator depends on the type of moderator and the type of reactor. In water-cooled and moderated reactors, the reactivity effects are initiated by changes in density which affect the slowing-down power and the absorption. Boiling-water reactors are operated at a relatively constant pressure and saturated conditions, which correspond to a relatively constant temperature. Density changes due to temperature are small. The steam production varies with power level so that density variation by voids is appreciable. In a pressurized-water reactor, the coolant in the core is subcooled and voids are suppressed. Coolant temperature may be varied with power, which would cause density changes as a function of temperature.

A reactivity change from the moderator based on the density change also occurs if there is dissolved neutron absorber (chemical shim) in the moderator. A density decrease causes less of the absorber to be present in a given volume and a decrease in neutron absorption with increasing temperature.

The reactivity effect of voids and of temperature change of the water, therefore, is to oppose a change in power while the reactivity effect, due to temperature change, of dissolved chemical shim is to aid a change in power. Since it is desirable (for safety reasons) to have a negative moderator temperature coefficient of reactivity, that is, a coefficient that with decreasing moderator density acts to retard the fission process, the amount of dissolved chemical shim is usually limited to that which will do no more than reduce the temperature coefficient of reactivity to zero.

Another reactivity effect is produced by the buildup of fission products. Some of these are neutron absorbers and will act as a retardant of the fission process by removing active neutrons. Two of the strongest absorbers are xenon, ^{135}Xe, and samarium, ^{149}Sm, whose absorption cross sections are 3×10^6 barns and 5×10^4 barns, respectively. ^{135}Xe is produced directly as a fission fragment (fission yield 0.3%), and in the decay chain of the fission fragment ^{135}Te (fission yield 5.6%) is

$$^{135}\text{Te} \xrightarrow[1\text{ min}]{} {^{135}\text{I}} \xrightarrow[6.7\text{ h}]{} {^{135}\text{Xe}} \xrightarrow[9.2\text{ h}]{} {^{135}\text{Cs}} \xrightarrow[2.1 \times 10^6\text{ yr}]{} {^{135}\text{Ba}}$$

On neutron absorption, ^{135}Xe is converted to ^{136}Xe, which is stable and has a low neutron cross section. ^{135}Xe, therefore, has two modes of production and two modes of elimination. This can be represented by the equation

$$\frac{dX}{dt} = (\lambda_x \Sigma_f - \sigma_x X)\phi + \lambda_x I - \lambda_x X$$
$$\frac{dI}{dt} = \gamma_I \Sigma_f \phi - \lambda_1 I \qquad (5\text{-}18)$$

where X = number of atoms of ^{135}Xe present per cubic centimeter (cm^3) at any time t
γ_x = fractional yield of xenon as a direct fission product
σ_x = microscopic thermal-neutron absorption cross section of ^{135}Xe
ϕ = thermal neutron flux
λ_1 = decay constant of ^{135}I
I = number of atoms of ^{135}I present per cm^3 at any time t
λ_x = decay constant of ^{135}Xe
Σ_f = macroscopic fission cross section of fuel in reactor
γ_1 = fractional yield of ^{135}I from direct fission process

The concentration reaches an equilibrium value during steady-state operation of the reactor but undergoes transients as the power changes. Of special concern to reactor regulation is the variation that occurs when the core is made subcritical following a power history. With the resulting large decrease in Xe removal by neutron absorption, the concentration of Xe increases because the difference in the ^{135}I decay and the ^{135}Xe decay allows a buildup. The peak concentration of ^{135}Xe is proportional to the preshutdown power level, as shown in Fig. 5-17. This absorbing effect must be overridden by control rods or elements if start-up is to occur.

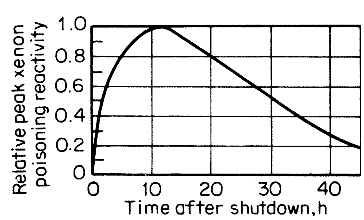


FIGURE 5-17 Relative peak xenon-poisoning reactivity.

Thermal Reaction. The heat from the core is generated in the fuel material as a result of the fissions initiated by the impinging neutrons. Although the center of a fuel element is subject to self-shielding, the fuel-element radial temperature distribution can be obtained by assuming a uniform volumetric heat source in a conduction problem. The gradient from the centerline of the fuel element through the gas gap and cladding to the coolant might be as shown in Fig. 5-18*a*. The power distributions (Fig. 5-18*b* and *c*) axially and radially in the core also vary, generally being highest in the center and decreasing toward the top and bottom and outside.

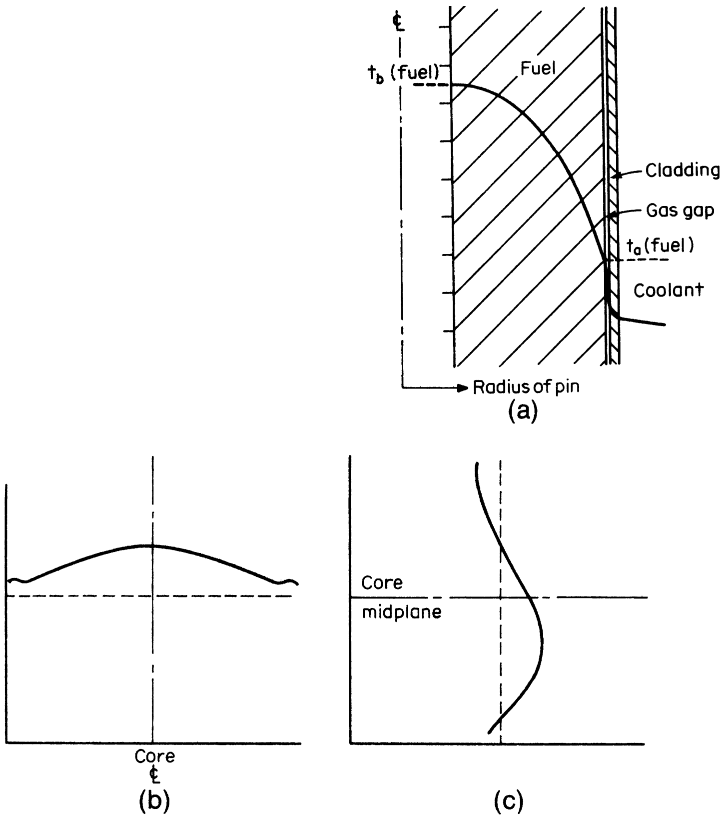


FIGURE 5-18 (*a*) Temperature distribution in fuel rod; (*b*) radial power profile; (*c*) axial power profile.

Control elements are adjusted to provide some degree of flattening. This makes a higher average power possible from the reactor for a given peak power.

For instrumentation and control purposes, the fuel may be considered to be mathematically represented by a series of time lags. These lags are important because the protection and control are strongly dependent on the time involved in transferring the heat out of the fuel. Since the elements are very long compared with their radius, conduction of heat in the axial direction can be neglected. In cylindrical coordinates, the Laplace equation for heat flow (neglecting the Z direction) is

$$\frac{\partial^2 T}{\partial \gamma^2} + \frac{1}{\gamma}\frac{\partial T}{\partial \gamma} + \frac{1}{\gamma^2}\frac{\partial^2 T}{\partial \phi^2} = \frac{\rho C}{k}\frac{\partial T}{\partial t} \tag{5-19}$$

where T = temperature, °C
t = time, s
k = thermal conductivity, W/(m)(°C)
C = specific heat, J/(kg)(°C)
ρ = density, kg/m^3
γ and ϕ = cylindrical coordinates, m

Solving this equation and assuming that the heat capacity of the cladding and the thermal resistance from the surface of the fuel pellet to the cooling channel can be neglected, gives the fuel-element transfer function

$$G(s) = (1 - \gamma) \sum_{\eta=1}^{4} \frac{F_n}{1 + \tau_n S} + \gamma \tag{5-20}$$

where $G(s)$ = fuel transfer function
γ = fraction of heat produced by photons
F_n = gain of nth term
τ_n = time constant of the nth term

Application of Performance Equations. The equations of performance for the various portions of the NSSS may be used for mathematical analyses or for development of simulations. Simulation of an NSSS is frequently desired. Such simulation enables (1) evaluation of system performance to assure that plant performance requirements are met, (2) performance evaluation of monitoring or control equipment, and (3) analyses of protection action to assure protective-system adequacy.

The simulation may be performed with an analog, a digital, or a hybrid computer; selection of the appropriate method should be based on the equipment to be simulated and the objectives of the tests to be performed.

The level of detail of the simulation also varies with the objectives of the test. For instance, modeling of the reactor core may be one node for gross thermal input to another component or be multinode to evaluate the performance of in-core instrumentation; 1-, 4-, 7-, and 38-node models are examples of core simulations that have been found to be useful.

Since commercial power-reactor systems are large and are oriented to power production, it is inconvenient to encumber them with apparatus and operations directed to obtaining their time-response characteristics, which interferes with normal operation. It has been found that the discrete nature of neutrons and the statistical nature of the fission process give rise to random fluctuations in neutron population or "reactor noise." The production, absorption, and leakage of neutrons can be considered analogous to the random flow, from emitter to collector, of electrons in a diode. Using power-spectral-density techniques, the reactor transfer function can be developed from an analysis of reactor noise.

5.2.13 Control Systems

Fission Regulation. The fission process is regulated by the absorption, in a controlled manner, of some of the neutrons which cause fission. The controlled absorption may be provided by the control rods or control elements which are mechanically inserted into, or withdrawn from, the reactor core or by absorber material dissolved within the reactor coolant. Because steam is produced directly from the coolant in a BWR, dissolved poisons are usually not used in normal operation of the BWR. In a gas-cooled reactor where the coolant gas is either helium or carbon dioxide, it is not feasible to provide dissolved absorbers in the coolant. Dissolved poisons, therefore, are used principally in PWRs. Where dissolved poisons or fixed poisons within the core are used, they are generally used for the purpose of accommodating core burnup. Changes in the reactivity for normal operation, initial start-up, planned shutdown, and restart are normally accomplished by control rods or control elements. Since the rate of change available from dissolved poisons or fixed burnable poisons is normally quite slow, the mechanically inserted control rods or control elements are also used for emergency shutdown.

Considerations involved in determining the characteristics of the control rods or elements are the amount of reactivity that has to be controlled, the position accuracy (which corresponds to the minimum increment of the reactivity) to be provided, the rate of reactivity that must be provided for operation, and the reliability of components. The reactivity requirements of a nuclear system are based on the planned rate of fuel depletion, the fission-product buildup, including the principal poisons xenon and samarium, inherent reactivity effects such as temperature or void changes, and the control range necessary for maneuvering to which the plant may be subjected.

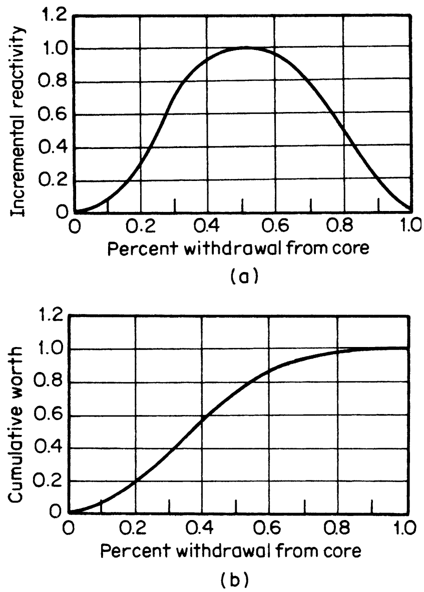


FIGURE 5-19 (*a*) Incremental reactivity versus core withdrawal; (*b*) cumulative worth versus core withdrawal.

The control rods or elements are also used to flatten the power developed radially and axially within the reactor to raise the average power and increase the output from the system. To accomplish this purpose, the rods or elements are normally assigned to groups or banks which are operated together. The capability of a rod or element to perform its neutron absorption or the worth of the control unit is dependent on the position of the element within the core. As the element progresses from the bottom, its worth increases from a low value to a peak and then decreases again to a low value as it is withdrawn from the top of the core. The typical incremental worth and the cumulative worth of an assembly are shown in Fig. 5-19. (Individual worths are affected by core power distribution.)

Emergency Shutdown. In order to provide an emergency shutdown capability, the control rods or elements are normally provided with a fast-insertion capability. This characteristic, also referred to as scram, is provided for control drives mounted on top of the reactor by a delatching capability with the rods or elements free-falling into the reactor core. With rods or elements that are mounted on the bottom of the reactor, a capability is provided to drive the rods rapidly up to their full insertion position. The latter mechanism is typically hydraulically actuated. Since the position of the control rods or elements is an important information in determining core power distribution and it represents a knowledge of the rate at which the fission process is proceeding, it is desirable to provide indication of the position of the control elements and to provide input from the rod-position sensors to core-power-distribution calculations. Position indication can be provided by analog meters, digital indicators, analog presentations on a cathode-ray tube, position logged by a digital computer with printout, or other types of display. Because of the importance of the position of the control element in reactor regulation and reactor shutdown, two systems of indication for the control elements are normally provided.

Controlling Fluid Processes. The primary system of most U.S. reactors involves either light water or helium as a coolant. Other fluid systems are provided to supply makeup, effect cleanup, process waste, etc. Conventional instrumentation is used to monitor these variables, and control signals in accordance with preselected control program are applied to appropriate actuator elements.

The special requirements of nuclear systems for continuity of cooling may require that extra care be used in the application of control equipment to assure high reliability. Establishment of set points and alarm points should also be done with due care.

Protection Systems. The high specific power of nuclear reactor sources coupled with the potential for the release of radioactivity, which might be a hazard to human beings, requires that additional systems be provided for regulation. In addition to the instrumentation provided for the control of the fission and fluid processes, systems are specifically supplied to initiate protective action in the event that preselected limits are exceeded. The relation of protection systems to other instrumentation and control systems is shown in Fig. 5-20.

Basically, a protection system must provide a functional capability to initiate action in the event of a design-basis event which, if unchecked, could lead to unacceptable consequences. Because of this requirement, the protection system must operate when required and must operate correctly. The development of functional and reliability requirements is very important in this arrangement.

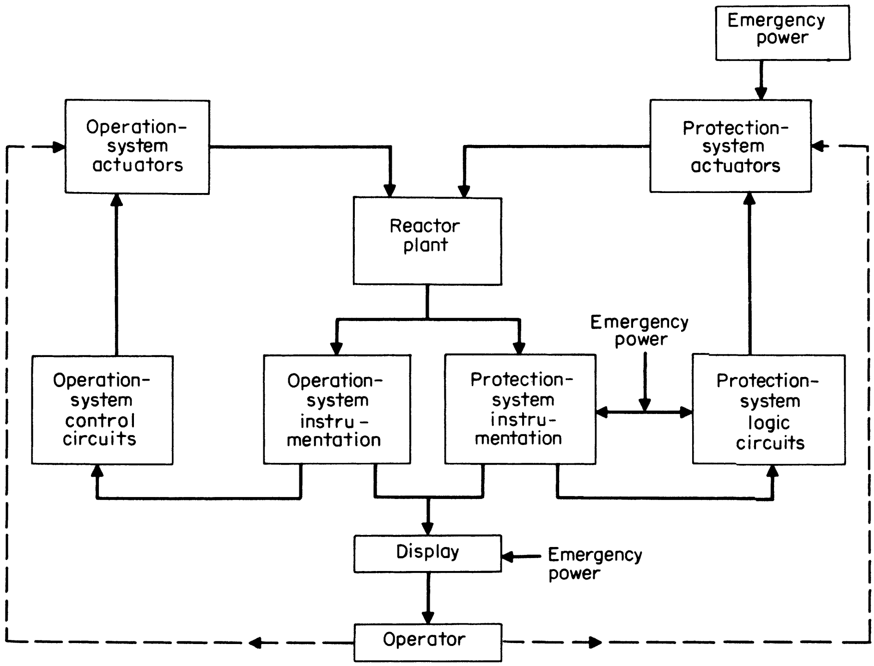


FIGURE 5-20 Relation of protection system to other instrumentation and control system.

The consequence of the greatest concern associated with a nuclear system is the release of radioactivity. The nuclear fuel itself has multiple barriers between the fuel material and the outside boundaries for the public. These barriers are the fuel matrix, the fuel jacket or cladding, the primary system, and the containment structure. For any given condition, it is the duty of the protection systems to prevent a situation which would lead to an excessive release of radioactivity from occurring past a given boundary.

The protection systems therefore include the reactor shutdown system and other systems that effect the containment of the radioactivity such as emergency core cooling, containment isolation, containment pressure reduction, emergency power sources, and air filtration. A protection system itself includes the instruments, logic systems, actuators, protective interlocks, and mechanisms which carry out the necessary functions. That system which effects reactor shutdown, normally referred to as the reactor protection system, includes all electrical and mechanical devices and circuits used to initiate a reduction of the fission process below criticality. The engineered safety features or engineered safeguards systems include everything else associated with protection except the reactor protection system.

In order to accomplish a high availability, that is, the capability to act when needed, it is necessary to provide multiple channels to perform the same action. The probability of any one channel not being able to act at a given time, therefore, is offset by having other channels capable of performing the action. In order to avoid spurious action, that is, initiation of a protection when none is required, it is usually the practice to provide a logic arrangement and initiate action only when there is a coincidence of two or more channels. The consequences of spurious action, for example, loss of a power source, in a large system can also be rather severe in terms of impact not only at the plant site but also in the area of public usage of the plant output, and such false action should therefore be precluded. Conventional protection systems have used a logic arrangement involving three or four channels and requiring the coincidence of two of three or two of four in order to initiate action.

In the selection of the plant variables that are sensed by the protection system, it is usual to evaluate those plant conditions which could occur as a result of some event. In some cases, such as those involved with protection of the clad integrity, the variables cannot be measured directly and an inferred or computed variable must be used to initiate the protective action. Such plant variables used to evaluate the fuel design limits are the departure from nucleate boiling (DNB) ratio in the case of PWRs and minimum critical heat-flux ratio (MCHFR) in the case of BWRs. In order to maintain these critical values below safety limits, it is necessary to monitor the observable parameters which affect DNB or MCHFR such as thermal power, coolant flow, coolant temperature, coolant pressure, and core power distribution (from the nuclear instrumentation and control-element-position sensors). These values are translated into systems relating to desired protection, and reactor shutdown is initiated if the measured variables approach boundaries of regions established by these systems. The earlier nuclear plants utilized parameters directly and the shutdown values or limits were set on the basis of calculated values relating to a set of curves. Current nuclear plants involve, to a certain degree, the use of online digital calculators permitting the protection system to provide continual computations of the relation of the variables. Shutdown limits are initiated as a function of the instantaneous value of the measured variables.

In order to avoid the possibility of disabling of the protection system by some common initiating mechanism, a principle of diversity of sensing and operation is suggested. The diversity relates to a different type of equipment or a different mode of operation to effect protection action from a given condition. Types of diversity that have been considered include equipment, functional, operational administrative, and design administrative diversity. Therefore, an evaluation should be made of the utilization of diversity in order to assure that (1) a definite objective may be obtained, (2) the additional complexity introduced by added equipment will not result in a degradation of the system, and (3) typical relations will be maintained between the primary action and the diverse action provided.

Nuclear Instrumentation. Since the fission process involves a neutron fluence where the fluence rate (flux) is proportional to power, it is essential to measure the fluence rate. Such measurement, for the large-core commercial reactors, involves special consideration, including the following:

Measurement over 10 to 13 orders of magnitude may be required.

There may be important spatial variations.

The measurement is of uncharged nuclear particles (neutrons).

The measurements have to be made in a background of substantial gamma radiation.

In addition to the monitoring of neutrons, a nuclear steam system involves the monitoring of radiation, principally gamma, from process lines and fuel.

Ex-Core Neutron Monitoring. Ex-core, or out-of-core, detectors are those which are located external to the reactor core and usually external to the primary pressure boundary. The fast-neutron flux leaking from the core provides a neutron flux spectrum for some distance beyond the vessel wall. This flux, proportional to a spatially average core power, becomes thermalized in the shielding, usually hydrogenous material surrounding the reactor.

The environment in which the detectors are located involves, typically, neutron fluxes from up to 10^{11} neutrons/(cm^2)(s), gamma dose rate up to 10^7 R/h, and temperatures to 200°C.

The detectors used are devices which produce a current pulse when subjected to the passage of a nuclear emission. The incident radiation drops some or all of its energy within the detector, causing ions to be produced. The ions produced are attracted to electrodes, within the detector, by the effect of voltage across the electrodes. The number of ions collected is a function of applied voltage. The chambers are filled with a gas selected to enhance the performance for a particular type of operation. The chambers are operated with voltages between the electrodes to effect the collection of charged particles as pulses or current (continuous pulses). Operation for ex-core ion-chamber detectors is normally in the flat or plateau region. The "knee" at the left of the curve moves to the right with increasing ambient radiation, so the selected voltage operating point must be sufficiently high so as to remain on the plateau for all conditions.

The chambers can be made sensitive to neutrons by coating the electrodes with a film of material containing an element with which neutrons interact, such as ^{235}U or ^{10}B. Enrichment of the isotope can be selected to effect desired performance in the range of neutron flux. In addition to pulses from incident neutrons, there are pulses from other incident ionizing radiation such as gamma rays. The contribution from gamma rays can be countered or compensated for by supplying two identical volumes, one with neutron-sensitive coating and one without, and subtracting their output. Since the neutron fluence rate and the gamma level do not increase together, that is, maintain the same ratio, this compensation can be completely canceled only at a given power level. Over the range of operation, about 97% to 98% compensation can be effected. The compensated ion chambers (CIC) can be used in a fluence rate about two decades below that of the uncompensated ion chambers (UIC). Ion chambers are applied for both in-core and ex-core monitoring.

Low neutron fluence rates produce currents too low to be accurately measured with an ion chamber. Proportional counters operating in a pulse-counting mode are conventionally used in this application. These detectors are filled with a gas, such as BF_3, which interacts with incident neutrons to produce ionized particles. Gas amplification is used to increase the output for a given event. Voltage applied is typically in the center of the proportional range. A well-regulated voltage supply is necessary.

Ion chambers using a sensitive coating involving ^{235}U absorb neutrons and undergo fission which generates the ions. Because of the substantial energy imparted to the ions by the fission process, these detectors are used satisfactorily for both current and pulse generation. Operation over 10 decades of neutron fluence rate may be satisfactorily achieved.

Reliability. Selection and procurement of the equipment must include consideration of minimum maintenance and low-failure-rate characteristics. The steps taken to achieve quality levels include good design practices, quality control, qualification testing, calibration, and system testing.

Independence. The equipment is to be installed so that independence of redundant channels is preserved. This requires that (1) components and circuits be electrically separated to prevent the propagation of electrical faults, (2) components and circuits be physically protected from destructive factors such as missiles and water or steam jets, and (3) steps be taken to avoid loss of protective action in the event of common-mode events such as fire or high temperature.

Signal Validation. The equipment design and arrangement is to be such that there are means of verifying that the signal represents the actual condition of the variable monitored. Such verification may include

1. Calibration
2. Cross checking between channels
3. Introducing and measuring known perturbation in the variable

Maintenance. In order to assure high system availability, redundant parts of the systems must be both repairable and adjustable. Special consideration must be given to access, bypassing, removal of modules, and calibration.

Information Readout. The system readouts are to be designed to provide operators with accurate, complete, and timely information. Consideration must be given to sequence and trend indication and to indication of related conditions.

Emergency Power Systems. Because of the need for protective action to be available at all times when the reactor is operating and the need for continued cooling and monitoring when the reactor is shut down, systems must be provided to assure high availability of electric power.

Primary Coolant Circulators. The largest single plant load is the drives for primary coolant circulation. Since it is important to maintain coolant circulation and since these drives are generally too large to be supplied by engine-driven sources, provisions should be made to supply the coolant circulator drives from two or more sources. Frequently, arrangements are made for the main generator to supply two or more power lines. Provisions in the switchyard enable the plant distribution system to be supplied from the plant generator or from one or more of the outside lines.

In spite of possible connection of plant loads to multiple external power sources, it is possible to lose all external lines, for instance, by a tornado. In this event, a local source of power to supply critical ac loads is required. For these purposes, engine (diesel)-driven generators are usually used. Credit can sometimes be taken for local hydro generators or gas-turbine generators if these sources can meet the requirements. These power systems must be designed so that they provide power to the station following a design-basis event. An ac power system (generation and distribution), a dc power

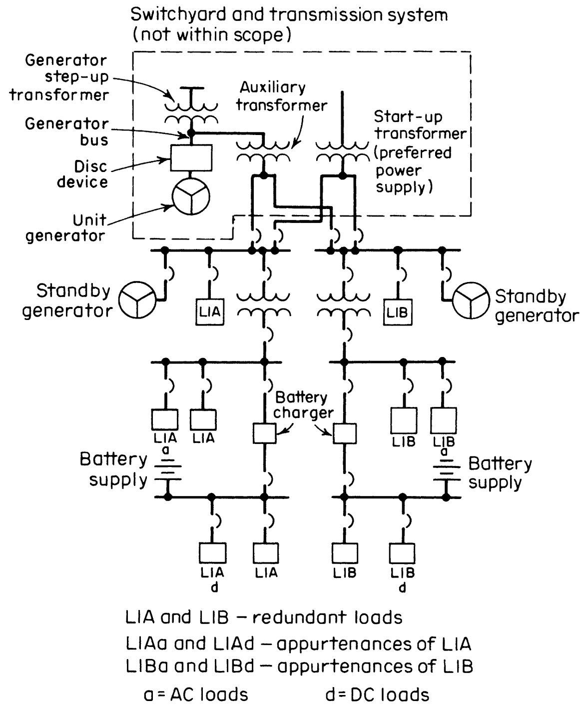


FIGURE 5-21 Class 1E power system for single unit.

system, and a vital instrumentation and control power system are provided. An example of a safety-grade power system is shown in Fig. 5-21.

In the ac system, each of the redundant load groups must have access to both a preferred and a standby power supply. The units of the standby supply must have sufficient independence from the preferred supply and from one another to preclude a common failure mode. Load assignment must be such that the safety actions of each group are redundant and independent. Protective devices must be provided to limit the degradation of the system and maintain the power quality (voltage and frequency) within acceptable limits. Following a demand for the standby power supply, it must be available within a time consistent with the requirements of the engineered safeguards features and the shutdown systems.

In the dc system, batteries, distribution equipment, and load groups are arranged to supply critical dc loads and switching and control power. Redundant load groups, and corresponding battery sections, must be sufficiently independent to preclude common failure modes. Each of the redundant load groups must have access to one or more battery chargers; the batteries are to be kept charged. The battery supplies must be sized to be able to start and operate their assigned loads in the expected loading sequence for a length of time commensurate with the protection provided.

Battery chargers supplying the redundant load groups must have sufficient capacity to restore the battery from its design minimum charge to its fully charged state while supplying normal and postaccident loads. Each charger supply must have a disconnecting device in its ac feeder and one in its dc output line. The dc system must be equipped with surveillance equipment to monitor its status and to indicate actions.

The vital instrument system is provided to power the instrumentation needed for reactor protection and engineered safety features. Since there may be considerable variation in the instrumentation in various plants, the vital system may be required to supply ac or dc or both. To preserve freedom from common-mode failure, the vital supply must be divided into redundant and independent systems with adequate status indication. Provisions for testing, adjustment, and repair should be included in the parts of the emergency power systems to improve reliability and availability.

5.2.14 Radioactive Waste Disposal

There are two broad categories of radioactive wastes, low-level wastes (LLWs) and high-level wastes (HLWs). Nuclear power plants produce significant quantities of both. (Other important sources include defense [military] labs and various commercial/university/hospital labs using radioisotopes, research reactors, etc.) Their safe and economic handling represents a key objective if nuclear power is to remain competitive in the future. Issues related to radioactive wastes are exceedingly complex, involving both intense political-societal-legal interactions and technology. Here we will briefly discuss various technical issues, but caution the reader that sociopolitical factors are, in many cases to date, overriding.

The principal Nuclear Regulatory Commission (NRC) regulation for LLWs is 10 CFR 61. It defines three classes of radioactive waste—A, B, and C—depending both on the isotope's half-life and on the specific activity in curies per cubic meter, class A representing the shortest half-life and less specific activity and class C being the highest. Ideally, the classification would be hazard-based, but for practical reasons and ease of enforcement, it is based largely on the process of the industry that generated the waste. 10 CFR 61 provides a quantitative evaluation for all radionuclides involved. All power-reactor fuel elements are automatically classified as HLWs while resin, sludges, contaminated clothing, materials, etc. are LLWs.

An LLW disposal facility can be licensed either by a state (if qualified by the NRC) or by the NRC itself for use by a commercial operator. The site is selected from appropriate candidate land when extensive facts about the geology, water flow patterns, and nearby population are known or developed. After 20 to 50 years, the site is closed, and ownership is transferred to the state or federal agency, which then continues to monitor the site for the period of institutional control (100 years). Then the license is terminated and no further maintenance is needed, but the design should have assured protection for a period of 500 years.

The disposal of HLWs is controlled by the Nuclear Regulatory Commission, following regulations 10 CFR 60. Some important provisions in this lengthy regulation are as follows:

The facility should not pose an unreasonable risk to the health and safety of the public; the allowed radiation dose limit is a small fraction of that from natural background.

A multiple barrier approach is to be used, including the waste form, the containers, and the site geology.

A thorough site characterization is required with features such as great geologic stability and slow water movement regarded as favorable.

The location should not have attractive resources, should be far from population centers, and under federal or state control.

HLWs are to be retrievable for up to 50 years.

Packaged wastes should be secure for at least 300 years; groundwater travel time to any source of public water should be at least 1000 years.

The annual release of radionuclides must be less than $10^{-3}\%$ (in radioactivity) present 1000 years after the repository is closed.

Predictions of safety must be made with conservative assumptions and by calculations that take account of uncertainties, using expert opinion.

In principle, prior to storage/disposal of radioactive isotopes, power reactor fuel elements could be reprocessed (e.g., by chemical solution followed by separation) to recycle fissile and some structural material so that only the radioactive fission products and activated substances are sent to disposal. However, by presidential decision, based on a concern that reprocessing might enable nuclear proliferation, the United States has not permitted reprocessing. (Other countries, e.g., France, have reprocessing plants.) The only actual reprocessing of commercial wastes done was by Nuclear Fuel Services, Inc., at West Valley, New York, in the period 1966 to 1972. This plant was shut down because it was uneconomical to operate, compared to the cost of producing fuel from uranium ore. Thus, since 1972, spent fuel has been accumulating at nuclear power plants.

Another distinction among radioactive wastes involves transuranic elements. Transuranic (TRU) wastes are those containing isotopes above uranium in the periodic table of chemical elements. They are the by-products of fuel assembly and weapons fabrication and of reprocessing operations. Their radioactivity level generally is low, but since they contain several long-lived isotopes, they must be managed separately. This classification includes isotopes with half-lives greater than 20 years with a total activity greater than 100 nanocuries per gram of waste material. Isotopes include plutonium-239, americium-241, americium-243, cirium-244, neptunium-237, and curium-245.

Transuranic wastes give off very little heat, and most of them can be handled by ordinary methods not requiring remote control. Since 1970, they have been placed in retrievable storage.

Another special category of wastes involved with the power reactor fuel cycle consists of mill tailings. These wastes are the residue from the physical and chemical processing of uranium ore to obtain uranium. They are stored and disposed of separately due to the enormous volume of material involved.

Another waste category must be considered along with "pure" radioactive materials. "Mixed wastes" are those containing both hazardous chemicals and radioactive substances. The disposal of hazardous wastes is regulated by the Environmental Protection Agency (EPA), while radioactive wastes are controlled by the NRC. Thus, it has been necessary to establish consistent dual rules by agreement between agencies. This overlap of agency control results in many complications, which have caused difficulty in handling of mixed wastes, and has led to most such waste being chemically processed to eliminate the chemical hazard, followed by disposal as LLRW.

What to do with spent fuel has been an important challenge for the United States since the early 1970s, and will continue to be well into the twenty-first century. It is estimated that about 80,000 metric tons of spent fuel elements will have to be stored by the year 2020, assuming, of course, that no reprocessing is done.

Several materials have been candidates for the host geologic medium for a disposal site. Regions containing deposits of rock salt, basalt, granite, tuff, and argillaceous materials (clay and shale) have been involved in several levels of investigation. The Yucca Mountain region in Nevada has been selected for detailed evaluation. A facility located deep beneath the surface of the tuff-type mountainous range would be employed. Tuff is a compacted and hardened ash that has come from volcanoes. Yucca Mountain is composed of a welded tuff, formed by a flow of ash into sheets, that has become solid through pressure from material above. The Yucca Mountain site would use the mined cavity method for storage (see the general concept illustrated in Fig. 5-22). A shaft would extend from the earth's surface down to a series of horizontal tunnels. Canisters containing spent fuel would be placed in holes drilled in the tunnel floor. Then the openings would be backfilled. In addition to the geological safety achieved, this method has the advantage of using conventional mining techniques. After a lengthy legal battle with the state of Nevada concerning the government's right to characterize the Yucca Mountain site, plus delays due to various citizens' protests about issues such as possible water seepage into the mine structure, the DOE recently resumed work to characterize and evaluate the site.

A large volume of wastes classified as high-level defense waste, including some mixed with toxic chemical liquids, is now stored in underground tanks and bins at three main government sites—Hanford, Idaho Falls, and Savannah River.

We see from the foregoing that there is one radioactive waste problem but three key disposal challenges: defense wastes; spent fuel; and low-level wastes.

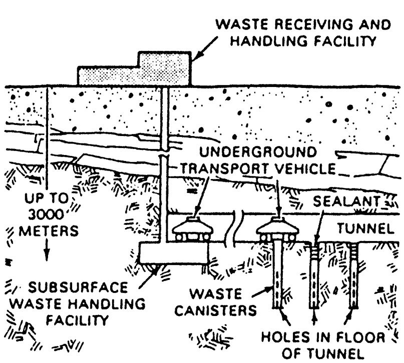


FIGURE 5-22 Emplacement of waste canisters in a mined cavity.

Modern reactors use uranium that has a higher percentage of ^{235}U (3%) than is found in nature (0.7%). The enriched fuel is typically in the form of uranium dioxide (UO_2) as small pellets about $^3/_{16}$ in in diameter and $^3/_8$ in long.

During the $1^1/_2$ to 3 years that a fuel element typically stays in the reactor, neutron irradiation has consumed some of the uranium and produced some new materials. The content of ^{235}U is reduced to about 1%, while ^{238}U has gone down from 97% to 94%. A new fissile isotope, ^{239}Pu, has been produced. It has a half-life of around 24,000 years, emitting an alpha particle of 5.1-MeV energy. As noted earlier, a small quantity of other transuranic elements are also produced. By successive neutron absorption ^{239}Pu becomes ^{240}Pu, fissile ^{241}Pu, ^{242}Pu, and short-lived ^{243}Pu. (Light-water reactors are sometimes called "converters" because they transform some of the ^{238}U into plutonium isotopes.) Figure 5-23 shows the before and after composition of spent fuel. Note that the spent fuel still has most of its original ^{238}U and a fairly high fissile fuel content.

The volume of a typical assembly is less than 7 ft^3 with about 60 assemblies removed per year from each reactor, a total of around 400 ft^3 of spent fuel must be handled. From the 100 reactors in the United States, the annual spent fuel volume would thus be around 40,000 ft^3, corresponding to less than 1-ft depth in a standard football field.

The actual amount of radioactive materials in the spent fuel is considerably smaller than the total volume quoted above. If the fuel were reprocessed, these would be extracted, as would most of the plutonium isotopes. The uranium would be cleaned up in preparation for reuse. For each 1200-lb assembly, there would be only around 35 lb of fission product waste. For 60 assemblies discharged per reactor per year, this is about 1 ton. The actual volume would be only 3.4 ft^3, which is 18 in on a side. However, most wastes could not be stored or disposed of in such a concentrated form because its radioactivity would produce intense heat.

If, as at present, fuel is not reprocessed, two choices are open: store it indefinitely or dispose of it by burial or other isolation technique. These options do not seem optimum in view of the value of the fissile elements in the fuel elements and the reduction in volume that reprocessing could achieve. The decision to avoid storage was based on the Pu-239 content and nuclear proliferation concerns. However, trade-offs among options must be constantly reevaluated with changes in both technology and international politics.

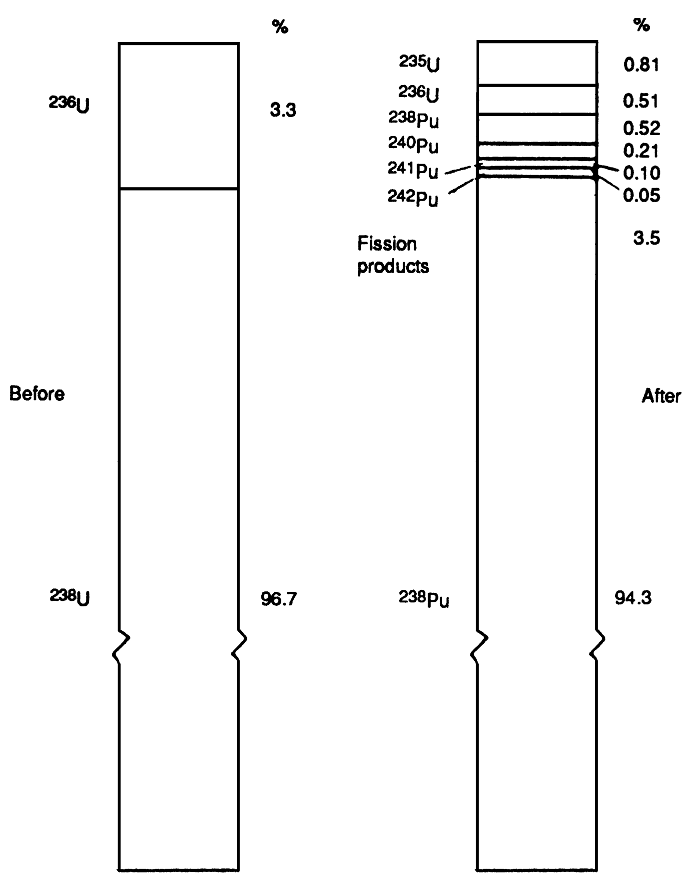


FIGURE 5-23 Composition (in percent) of fresh fuel and spent fuel. Uranium-235 is burned to form fission products and ^{236}U; ^{238}U is converted into plutonium.

5.2.15 Prior and Present Trends in Nuclear-Fueled Plant Development

The development of nuclear-fueled steam-electric plants underwent substantial change in the 1970s. At the beginning of the decade, orders for nuclear-fueled plants were increasing to a peak of 38 per year. Following the oil crisis of 1973 to 1974, changes in the economy began to affect the cost of, and consequently the demand for, electric power. Opposition to the use of nuclear energy for electric power production increased; litigation was frequently employed. Near the end of the decade, sociopolitical aspects of nuclear-fueled plants became as involved and time-consuming as the technical aspects. In order to participate effectively in the design, construction, and operation of nuclear-fueled plants, one must be familiar with the energy perspective; the concerns about the use of nuclear energy; and the functions of advocates, intervenors, and regulators.

With the maturity of the nuclear-fueled plants, more emphasis was placed on project management (Pederson 1978). Siting of the plants became a major task (Winter and Conner 1978). Because of the reduced demand for electric power, the increased cost of money, and the difficulty of resolving the objections raised, orders for nuclear-fueled plants began to decrease sharply after the middle of the decade. Some orders were canceled. Then in March 1979, a major accident occurred at the Three Mile Island plant, causing serious damage to the plant. This event raised questions about the operation of nuclear-fueled plants and a review of the value of nuclear energy (Rubenstein 1979). At the end of the decade, orders for new nuclear-fueled plants had been reduced to zero and a sizable number of plant orders had been canceled.

The beginning of the decade of the 1980s saw reinforcement of the need for commercial use of nuclear energy (Greenhalgh 1980), but also heralded changes in the safety, control, and maintenance systems. In the electrical area, the most notable changes were the redesign of control rooms and stations and the increased use of computers in more sophisticated safety systems (Hanes et al. 1982). The study of incidents and malfunctions by means of computers has provided another means to inform and guide operators and to evaluate possible trouble spots (Kaplan 1983). The availability and capability of the microprocessor has provided new ways to improve the safety and performance of plant instrumentation, control, and safety systems.

With fewer new nuclear plants being built worldwide than originally anticipated, much attention has been on methods to achieve "life extension" of present plants (retrofitting to allow operation beyond the traditional 20-year life cited for power plants). At the same time, procedures for decontamination and decommissioning of plants being shut down are being refined. The NRC is simultaneously developing streamlined procedures for licensing new plants, with the anticipation that utilities may turn to nuclear energy in the future in the form of the new passive-safe type reactors. This effort, the deregulation of the utility industry in the United States, plus the possible emphasis on nuclear energy as a way to meet goals for reduction of CO_2 greenhouse gases (Schmidt 1998), could have a profound effect on the evolution of the nuclear industry.

There has been a growing belief in recent years that a 'rebirth" of nuclear energy has begun. This has been driven by the rapid increase in oil process coupled with a desire by countries like the U.S. to achieve energy independence, while future energy needs will be met by a combination of conservation plus use of a wide range of energy sources (solar, wind, bio renewable energy sources). There is a growing opinion that nuclear is needed to fill the future needs for central plant energy production. The new "Generation-IV" reactor concepts being developed by DOE offer many attractive features to fit into this need, including competitive costs, passive safety, improved efficiency, reduced maintenance down time and costs, and they also eliminate greenhouse gas emission. The main obstacle remaining is gaining of public acceptance of radioactive waste disposal techniques.

BIBLIOGRAPHY

Cook, C. E., 1980. *Nuclear Power and Legal Advocacy*, Lexington, Mass., Lexington Books.

Greenhalgh, G., 1980. *The Necessity for Nuclear Power*, London, Graham and Trotman.

Hanes, L. F., O'Brien, J. F., and DiSalvo, R., 1982. Control Room Designs; Lessons from TMI, *IEEE Spectrum*, vol. 19, no. 6, pp. 46–52.

Kaplan, G., 1983. Nuclear Power Plant Malfunction Analysis, *IEEE Spectrum*, vol. 20, no. 6, pp. 53–58.

Meghreblian, R. V., and Holmes, D. K., 1960. *Reactor Analysis*, New York, McGraw-Hill.

Murray, R. L., 1989. *Understanding Radioactive Waste*, 3d ed., Columbus, Oh., Battelle Press.

Pederson, E. S., 1978. *Nuclear Power Project Management*, Ann Arbor, Mich., Ann Arbor Science.

Rubenstein, E., 1979. Three Mile Island and the Future of Nuclear Power, *IEEE Spectrum*, vol. 16, no. 11, pp. 30–111.

Schmidt, Karen, 1998. Coming to Grips with the World's Greenhouse Gases, *Science*, July, vol. 281, p. 504.

Warnock, D., and Bossong, K., 1979. *Nuclear Power and Civil Liberties; Can We Have Both*, Washington D.C., Citizens Energy Project.

Winter, J. V., and Conner, D. A., 1978. *Power Plant Siting*, New York, Van Nostrand.

5.3 NUCLEAR POWER FOR THE FUTURE

By GEORGE H. MILEY

Development in nuclear power for the future is following three main paths: the extension of present reactor designs to create advanced concepts with passive safety features and improved economics; the development of breeder reactors to extend the supply of fissionable fuel; and the development of a new nuclear energy source, fusion. Each of these paths is considered in turn in the following sections.

5.3.1 Advanced Concepts with Passive Safety Features

Recent progress in advanced nuclear power development reveals a high potential for nuclear reactor systems that are smaller and easier to operate than the present generation. Passive, or intrinsic, characteristics are applied to ensure continued cooling of the fuel and its containment systems even in the advent of a major breakdown of the normal cooling/control functions. This substantially reduces the chance of a severe accident. The reactor design concepts that are emerging are simpler, more rugged, have a longer lifespan, and place less burden on equipment and operating personnel. Modular design concepts and design standardization are envisioned to reduce construction time and engineering costs.

Common Development Goals. The primary thrust in U.S. advanced reactor development is to incorporate design improvements to achieve five primary objectives:

Assured safety with features to minimize negative consequences of human error, especially a reduction in the probability of severe core damage.

A significantly simpler design, with increased safety and performance margin in key operational parameters.

High reliability throughout a lifetime of about 60 years.

Reduction in costs to meet the economic competition.

A standardized design which is predictably licensable.

Common generic technical features (passive stability, simplification, ruggedness, ease of operation, and modularity) are being developed to respond to these goals for each of the principal nuclear power systems—the light-water reactor (LWR), the liquid-metal reactor (LMR), and the high-temperature gas-cooled reactor (HTGR). These features, coupled with standardization and assurance that the plant is licensable, should ensure future economic competitiveness.

Light-Water Reactor with Passive Safety Factors. An effort to develop a passively stable LWR in the United States is jointly sponsored by the Electric Power Research Institute (EPRI) and the Department of Energy (DOE) with substantial contributions from several major U.S. suppliers.

Conceptual designs have been developed for passive versions of both a boiling water reactor (BWR) and a pressurized water reactor (PWR). A 600-MWe unit output is proposed to provide the utilities the option of a smaller nuclear power plant and to make it easier and less costly to incorporate passive cooling features.

PWR with Passive Safety Features. An advanced PWR, called AP-600 uses proven technology: a UO_2-fueled core and field-proven plant components. The burden on systems has been reduced by increasing design margins through reductions in coolant temperature, flow rate, and core power density and by selecting higher-quality materials and more robust components.

Passive cooling is provided by a passive emergency core cooling system (ECCS) and a passive containment cooling system. The ECCS consists of a combination of cooling water sources: gravity drain of water and water ejected from two accumulator tanks under nitrogen pressure. Additionally, the system can be depressurized to permit an even larger amount of water to enter the system. If a feedwater accident disables the steam generators, core decay heat is removed through a passive residual heat exchanger. This transfers core decay heat to the refueling water by natural circulation.

Containment integrity is ensured by cooling the containment shell by evaporating water that is gravity-fed from a large storage tank; heat is removed by a natural-circulation air system. Only automatic valve operations are required to provide emergency core cooling and containment cooling after a major energy release.

It is estimated that the use of a modular design for the AP-600 will reduce construction time to 3 or 4 years. The simplified design combined with the shorter construction time should counter the loss of economy of scale that previously favored larger plants.

The Passively Stable BWR. A passively stable BWR design uses fully proven components and systems operating at reduced burdens: lower power density and increased thermal margin. The reactor operates at full power under natural circulation, which eliminates the need for recirculation pumps while reducing the amount of piping, valves, and controls used in present BWRs.

A passive containment cooling system (PCCS), which consists of three condenser units, provides the capability for removal of the decay heat for a period of 3 days. In the event of a LOCA, the steam discharged from the break and noncondensibles from the dry well flow through the condensers to the wet well. The steam is condensed and the condensate is returned to the gravity-driven cooling system pools. Isolation condensers are used to remove decay heat during reactor isolation events. The condensed steam from the reactor is returned to the reactor vessel, maintaining vessel inventory and limiting the pressure increase without the need to open safety/relief valves. Both the isolation condensers and the passive cooling condensers are located in a large interconnected pool of water on top of the containment building.

The Advanced Liquid-Metal Reactor (LMR). The LMR development work is focused on a modular, passively stable reactor concept called S-PRISM (stable power reactor inherently safe module).

The principal parameters of PRISM are thermal power, 4000 MWt, from four nuclear modules; electric output, 1500 MWe, from two turbine generators; net efficiency, 37.5%; steam conditions, 864°F/2400 lb/in^2 (absolute); exit sodium temperature, 950°F; core power density, 199 W/cm^3; and equilibrium fuel burnup, 150 MW · day/kg. The two reactor modules in each power block provide heated sodium to a single sodium-to-water heat exchanger that generates steam for a single 750-MWe turbine generator. A commercial S-PRISM plant is envisioned to consist of a series of three such 465-MWe power packs, each of which would be functionally independent of the other two.

The reactor vessel auxiliary cooling system (RVACS) provides for emergency core cooling after any incident that causes a loss of the normal emergency heat-conversion systems. The residual-heat removal path consists of radiant-heat transfer from the reactor vessel to the containment vessel where the heat is removed by natural circulation of air between the containment vessel and the biological shield. This combination of passive reactor stability and passive cooling provides assurance of residual-heat removal without operator action.

The reference fuel for the S-PRISM concept is a uranium-plutonium-zirconium alloy with plutonium concentrations of about 25%. An important feature potentially achievable with this fuel is use of a metallurgical processing method for the separation of uranium, plutonium, and the transuranic elements from the fission products. Thus, long-lived transuranics can be recycled in the LMR rather than sent to a disposal site. This approach is basic to the integral fast reactor (IFR) concept pioneered by the Argonne National Laboratory (ANL). The S-PRISM concept is also capable of using standard oxide fuel in the event that the development promise of the metal fuel is not realized.

ANL has applied its experience in metal fuels operation, fabrication, and reprocessing to develop the IFR concept, which envisions a colocated nuclear power plant, fuel fabrication, and fuel reprocessing center where S-PRISM could function as the power plant. Although such a collocated concept is not essential to S-PRISM, or vice versa, this would provide for greater proliferation resistance because plutonium-bearing materials would not have to be transported outside the security boundaries of the site. The burning of the long-lived actinides via the colocated plant would also significantly ease the waste storage problem.

Advanced Modular Gas-Cooled Reactor. The modular high-temperature gas-cooled reactor (MHTGR) has been the primary focus in advanced gas-cooled reactor development. Its principal parameters are thermal power, 1400 MWt, from four nuclear modules; electric output, 538 MWe from two turbine generators; net efficiency, 38.4%; steam conditions, 1005°F [2515 lb/in^2 (abs)]; core exit coolant temperature, 1268°F ; core power density, 5.9 W/cm^3; and equilibrium fuel burnup, 92.200 MW · day/ton.

The reactor core is composed of hexagonal blocks of graphite-fuel elements in an annular array. The fuel is in the form of coated particles of low-enriched uranium oxycarbide and thorium oxide. This fuel and the graphite moderator arrangement have been termed a "prismatic fuel." It provides an essential barrier to fission product release during an accident. Test data have shown that essentially no failure of the refractory coatings occurs if the fuel is maintained below 1800°C. Even if all active cooling systems were unavailable, decay heat is dissipated by conduction and radiation to the reactor cavity cooling system in the reactor enclosure. This limits the maximum fuel temperatures in accidents to about 600°C, well below the fuel failure temperature.

Tests at a German gas-cooled reactor demonstrated this passive cooling capability after a loss of all coolant flow with no intervention by the operator. The conceptual design of the MHTGR is presently under review by the U.S. Nuclear Regulatory Commission to assess licensability.

Other Advanced Reactor Design Projects. The major industrial countries outside the United States also have advanced nuclear power development programs, several of them larger than those in the United States. The largest programs are in France and the United Kingdom, where the LWR and LMR are under development; and in Germany and Japan, where all three types of reactors are studied. ASEA Brown Boveri in Sweden has proposed a passive LWR, called the process inherent ultimate safety (PIUS) reactor, a concept originating from studies of reactor systems suitable for central heating applications. PIUS is a 640-MWe PWR plant; its core is enclosed in a large prestressed concrete vessel. A fluidic valve is located at the bottom of the core. It introduces, through intrinsic thermal-hydraulic properties, emergency core cooling from the pool of water surrounding the reactor. The concept reduces active equipment beyond the level of U.S. passive designs discussed here.

The Canadian nuclear industry is continuing its development of the 600-MWe heavy-water-cooled reactor called CANDU. This reactor has achieved a superior performance record thus far. A 300-MWe CANDU is being developed to provide a smaller-size system for the utilities.

Conclusion. The advanced reactor approach to next-generation reactors will require a substantial development program to verify the safety and cost objectives of the advanced designs. This includes detailed design, licensing review, cost estimates, extensive testing of the LWR and SBWR passive-cooling features, and prototype operations of the modular LMR and HTGR advanced systems. While a variety of concepts have been proposed, they all involve systems that increase the use of passive systems (versus active equipment) to protect against accidents; that is, systems that reduce the dependence on rapid operator response to abnormal conditions, systems that are simple and rugged, and systems that perform reliably without requiring high performance of equipment.

5.3.2 Breeder Reactors

The breeder reactor is able to create more fissionable material than it consumes. A fast breeder is designed to have more nuclear reactions created by neutrons in the high-energy region than by neutrons in the thermal-energy region, as used by the light-water reactors. This design is accomplished by not moderating, or slowing down, the energy of the neutrons created by one fission reaction prior to subsequent captures of, or fissions by, the neutrons created. In the light-water reactors previously described, the water coolant absorbs energy from the neutrons, creating a condition (called the thermal-energy region) in which more neutrons fission the $^{235}U_{92}$ isotopes than are captured by the $^{238}U_{92}$ isotopes. The capture of a neutron by the $^{238}U_{92}$ isotopes results in a decay chain that produces an atom of plutonium (see reaction chain following). The plutonium can be used to fuel nuclear reactors. Liquid-metal fast-breeder reactors (LMFBRs) operate in the "fast," or high-energy region, where more of the neutrons are captured by fertile material than cause fission in fissionable material, and where the average number of neutrons created in fission reactions by fast neutrons is greater than the average number created by thermal neutrons. These additional neutrons permit the breeder to produce more fuel by the following reaction chain than is consumed by the fission reactions.

$$^{1}n_0 + {}^{238}U_{92} \longrightarrow {}^{239}U_{92} \xrightarrow[23\text{ min}]{\beta^-} {}^{239}Np_{93} \xrightarrow[2.3\text{ days}]{\beta^-} {}^{239}Pu_{94}$$

This same reaction occurs in a light-water reactor, but, without the greater number of neutrons and higher ratio of capture cross section to fission cross section, less fuel is created than is consumed. Much of this bred fissionable material is fissioned in situ. For the excess fissionable material in a breeder to be of use in other reactors, or for refueling itself, the spent fuel and blanket assemblies require reprocessing to separate the fissionable material from the fission products created by the nuclear fissions. These fission products, and the damage to the material of the fuel tubes, also limit the life, or time, breeder fuel can remain in the core to produce power. The fission products compete for capture of the neutrons and eventually "poison" the fission reactions, shutting down the chain reaction.

Thermal breeder reactors are possible using a $^{233}U_{92}$ fuel. This isotope results from the decay chain started from the fertile material thorium, $^{232}Th_{90}$.

$$^{1}n_0 + {}^{232}Th_{90} \longrightarrow {}^{233}Th_{90} \xrightarrow[23\text{ min}]{\beta^-} {}^{233}Pa_{91} \xrightarrow[27\text{ days}]{\beta^-} {}^{233}U_{92}$$

Experimental thermal breeder reactors have been operated using a seed-and-blanket core with water coolant or a molten-salt homogeneous core. The helium-cooled graphite-moderated reactor can also theoretically perform as a breeder.

The present breeder reactor designs use a liquid metal (sodium) to cool the core without moderating the energy level of the neutrons. This liquid metal becomes radioactive (15.5 h half-life) but is isolated from the condensate-feedwater-steam by a secondary, or intermediate, liquid-metal coolant circuit. Thus, an additional heat exchanger and pump are necessary with the attendant impact on plant cost and operating complexity.

Because all fuel resources are finite, breeding of fuels will eventually be necessary to sustain the present use of power with an increasing population and diminishing reserve of fuel. Crude oil and natural gas fuels will be depleted for practical purposes in 100 years. Coal resources will be depleted in about 450 years. Other concerns, such as the greenhouse effects, may limit the use of coal. Without breeding, uranium will be depleted in less than 100 years, as only 0.7% of natural uranium is the fissionable isotope $^{233}U_{92}$. Breeders will permit most of the uranium and thorium resources to be used as fuel. Breeders can supply our energy needs for thousands of years while avoiding production of greenhouse-effect gases. At a growth rate of 3% per year in power utilization, breeder technology needs to double the fuel every 23 years. Present experimental LMFBRs have demonstrated that *even shorter* doubling times are attainable.

Present Status. In 1983, funds for completing the Clinch River breeder reactor, a power-demonstration fast-breeder reactor, were canceled. Subsequently, the U.S. breeder program has regrouped to focus on the advanced LMR development program described earlier. However, this program was also terminated in the mid-1990s. Russia, France, Germany, Great Britain, and Japan also continued their research and development support of breeder reactors until recent events caused curtailment (but not cancellation) of their efforts.

Breeder Reactor Developments in Europe. Breeder reactor technology is being developed in Europe by combining national with cooperative ventures. Two prototype pool-type plants of 250-MW capacity each, Phenix in France and PFR in the United Kingdom, have been operated. A loop-type plant, SNR 300, was built in Germany but was canceled and terminated. A commercial-size plant, Superphenix, has been in operation in France but is in danger of being decommissioned.

These projects, done on a national basis in the early 1980s, resulted in designs known as Superphenix 2, CDFR, and SNR 2 in France, the United Kingdom, and Germany, respectively. Since 1984, these countries have jointly focused on one single design, the European fast reactor (EFR), which should incorporate the best features from the various national designs.

Breeder Reactor Development in Japan. The LMFBR is being developed in Japan as a national development project. The goal is to develop large oxide-fueled LMFBRs. Japan plans to bring these plants to commercialization by 2020 to 2030.

Joyo, a 100-MWt experimental fast reactor, has been operating successfully since 1977. Monju, a 280-MWe prototype FBR plant, began construction in October 1985, and achieved initial criticality in 1993. The plant was shut down, however, following a 1996 accident.

A demonstration FBR plant, now planned to follow operational experience with Monju, is expected to demonstrate the prospect for commercialization of the LMFBR. A preliminary design study of the demo FBR is in progress in order to verify the technical feasibility of a top-entry loop-type FBR. The plant assumed for this study is to produce between 600 and 800 MWe.

In summary, the Japanese FBR development program is quite aggressive. The Joyo and Monju reactors provide valuable operational experience. A variety of advanced concepts are under active consideration for the next-step demonstration reactor scheduled for operation in the early 2000 time frame so that competitive commercial plants can be introduced by 2020 to 2030.

Breeder Reactor Development in Other Countries. In addition to the U.S., European, and Japanese programs, active work on liquid-metal fast breeders is in progress in the former Soviet Union and India.

Fast-reactor development began in the former Soviet Union in the 1950s. Two reactors, BN-350 and BN-600, are in operation. Design and development of future fast reactor concepts have been initiated with improved safety and better economics as prime goals.

Because natural uranium availability in India is limited, a FBR system has been chosen for increasing the supply of fissile material to meet future energy needs. The fast-breeder test reactor (FBTR) was the first major step in this direction and is to be followed by the construction of a 500-MWe liquid-sodium-cooled fast-breeder reactor, called *prototype fast-breeder reactor* (PFBR). Design and construction of PFBR is expected to be the beginning of the commercial deployment of such reactors.

The FBTR attained its first criticality in October 1985 and began producing electricity in 1991. The first core is small, producing only 16 MWt. The next core to be introduced will be close to the reference design of about 35 MWt. FBTR has served the purpose of a pilot plant and has also turned out to be a good investment for developing trained workers. This experience will lead to the next significant step: the 500 MWe PFBR.

Projections for future energy requirements have driven the India breeder program. It is estimated that India will need over 3000 billion units of electricity by 2040. The goal for FBRs is defined by the need to supply at least 1000 to 2000 billion units so as to significantly reduce the use of coal for electrical production. An installed capacity of 200 to 400 million kW in fast breeder reactors would be required to meet this goal. Starting with an annual feed of 3000 kg of plutonium produced from 10 million kW of PHWR capacity, the realizable growth for different fuel options is given in Table 5-5.

Thorium is not introduced in the fuel cycle in any significant way. Thus, the fast reactor program would be based on Pu–U-238 cycle as long as depleted uranium is available as a by-product of the first-stage heavy-water program. Preference for depleted uranium is based on higher growth potential of a Pu–U-238 fuel cycle.

Table 5-5 shows that a ternary metal alloy comes closest to meeting India's needs when the fuel is designed to give a specific power of about 0.3 MWe/kg of fissile material and a breeding ratio of 1.4 to 1.5, corresponding to an annual net breeding gain of 316 kg/GWe. Advances reported for the

TABLE 5-5 Growth of Fast Breeder Installed Capacity

Characteristic	Fuel type (U + Pu) Mixed oxide	Mixed carbide/nitride	Metal
Specific power MWe/kg	0.25	0.30	0.30
Breeding gain, kg/GWe · year, allowing for processing losses	130	188	316
Installed capacity, GWe · year			
2010	7.8	10.4	11.6
2030	30.8	54.6	90.6
2050	53.2	150.2	430.4

Note: 3000 kg/year fissile material from 10000-MWe PHWRs; 3 years for recovery of bred material; process loss per GWe/year, 14 kg oxide and 28 kg carbide and metal.

development of the ternary metal alloy make the assumed figures appear realistic. Research and development for fuel development has therefore shifted from the mixed carbide to the ternary metal alloy.

The Indira Gandhi Center for Atomic Research and Nuclear Power Corporation of India Ltd. have created a new public sector company Bharitiya Nabhikiya Vidyut Nigam Ltd. to carry out the construction and operation of the India Department of Atomic Energy's first 500-MWe prototype fast breeder reactor.

5.4 NUCLEAR FUSION

By GEORGE H. MILEY and STEPHEN O. DEAN

The substantial, research effort now in progress aimed at controlling thermonuclear reactions to produce electric power at gigawatt levels is stimulated by the fact that the basic fuel consumed in fusion, deuterium, naturally occurs to the extent of one atom for every 6500 atoms of hydrogen, making ocean and lake water the basic fuel. This represents a virtually inexhaustible fuel source. Fusion liberates energy by combining two nuclei of light elements into one nucleus of a heavier element. The resulting mass is less than that of the fusing nuclei. Fission, results in splitting one atom of a heavy element into two atoms of lighter elements. The resulting mass, again, is less than that of the fissioning atoms. Each fusion reaction releases about 20 MeV of energy. Each fissioning reaction releases about 200 MeV of energy. (See also Sec. 5.4.)

Fusion will enable use of an extremely large fuel resource—the deuterium isotope, $^2D_1$, which exists as 1 part in 6500 parts of $^1H_1$ in natural hydrogen. The deuterium available from the ocean could fuel the power requirements of the world for billions of years—even until the Sun expands and envelops the Earth. However, research is still far from developing a commercial fusion reactor. Fission, solar energy and coal, other renewables, and nuclear fusion provide the means of avoiding severe shortages of electricity and other forms of power in the twenty-first century.

Many years will be required to develop a commercial fusion power plant. Current estimates for the introduction of commercial fusion plants are generally for sometime in the mid-twenty-first century. In fiscal year 2004, the Dept. of Energy outlay in support of civilian fusion research in the United States exceeded \$263 million. A larger budget, \$514 million, is provided by the National Nuclear Security Administration for inertial confinement fusion aimed at simulating nuclear weapons.

5.4.1 Fusion Reactions

The reactions of interest for fusion power occur when nuclei (ions) of elements of low atomic number are brought together in a plasma, at such temperature (i.e., velocity) as to give the nuclei sufficient speed to overcome the repulsive coulumbic barrier between them (and fuse), thereby transforming a part of their mass into kinetic energy. Because it is most reactive, the reaction generally favored for first-generation fusion plants involves deuterium, tritium, and lithium, in the so-called D-T-Li cycle.

This cycle involves two types of reactions. In the first, deuterium and tritium react to produce an alpha particle (helium-4 nucleus) and a high-energy neutron:

$$\mathrm{D} + \mathrm{T} \rightarrow {}^{4}\mathrm{He} + n$$

The helium-4 ion emerges with an energy of 3.5 MeV, and the neutron with 14.1 MeV.

The second reaction involves the production of tritium by lithium, which is contained in a "blanket" surrounding the plasma. Two neutron-induced reactions are involved in the production of tritium:

$$^{6}\mathrm{Li} + n \rightarrow \mathrm{T} + {}^{4}\mathrm{He}$$

and

$$^{7}\mathrm{Li} + n \rightarrow \mathrm{T} + {}^{4}\mathrm{He} + n \qquad (E_n \gtrsim 7\ \mathrm{MeV})$$

The tritium used in the D + T reaction is thereby regenerated by the lithium reactions in the "blanket" region external to the plasma. The power density produced by this reaction reaches a maximum when the kinetic temperature of the plasma is about 15 keV.

5.4.2 Advanced Fuels

A number of fusion reactions are possible at temperatures of potential interest for fusion (<1 MeV) that involve various light elements through boron. The principal advantages of seeking such "advanced fuels" include elimination of the need to breed tritium, reduced neutron fluxes (to reduce both materials damage and induced radioactivity in vessel and blanket structure), and an increased fraction of fusion energy carried by charged particles (making direct energy conversion attractive). Two general classes of fuels are of interest: deuterium-based and proton-based. Examples of the former include deuterium-deuterium (D-D) and deuterium-helium 3 (D-^{3}He) fusion while hydrogen- (proton)-boron 11 (p-^{11}B) is a frequently cited proton-based reaction. The deuterium-based reactions all involve neutron production to some degree due to D-D reactions whereas p-^{11}B represents a fairly ideal fuel since the primary product is helium ($p + {}^{11}\mathrm{B} \rightarrow 3\alpha$). The supply of hydrogen and boron 11 (from borax) is virtually as extensive as for deuterium; ^{3}He, however, must be bred, either via D-D reactions or through the decay of tritium (12-year half-life). Another potential source of plentiful ^{3}He that has gained considerable attention involves lunar mining. Bombardment of the lunar surface by the solar wind has resulted in the impregnation of the upper crust with a variety of gases, including ^{3}He.

Because of the relatively small fusion cross sections involved and the higher energies (plasma temperatures) required, all of the advanced fuels pose more stringent confinement requirements and offer lower power densities compared to D-T-Li fusion. Consequently, these fuels appear to be candidates for later-generation power plants. Still, the improved environmental compatibility they offer, combined with a significant reduction in neutron damage to structural materials, make this an important long range goal.

5.4.3 Power Production

The major part (about 80%) of the energy released by the D-T-Li cycle is carried by the high-energy neutrons. These, being chargeless, cannot interact with electric or magnetic fields, and hence cannot produce electric power directly. Rather, their power is converted to heat, which is extracted from the lithium blanket by thermal transfer. This heat (like that in fission- and fossil-fueled systems) must be converted to electricity by a thermal system, for example, one employing vapor (e.g., steam) or a working liquid (e.g., potassium).

The remaining 20% of the power generated involves charged alpha particles (helium-4 nuclei). This power remains in the plasma, thus sustaining its temperature. With a p-^{11}B cycle, the charged particle energy could be increased to nearly 100%, making direct conversion a more important avenue for energy extraction.

Energy Breakeven. In nuclear fission, the reaction is neutron-induced and self-sustaining when criticality is reached. In nuclear fusion, on the other hand, the strong mutual repulsion between the positively charged deuterium and tritium nuclei must be overcome to cause them to fuse. To do so, power must be supplied to start and heat the plasma to create reactions. The power derived from the fusion reactions must, of course, exceed the input power. Energy breakeven is the condition at which the fusion energy released just equals the energy input required to heat the plasma to the fusion temperature.

To achieve energy breakeven, the product $n\tau$ of the plasma density n and the time τ during which the plasma is confined must lie within a narrow range for a specific reaction. In the D + T reaction this product (i.e., $n\tau$) must be roughly 10^{14} ions/cm$^3 \cdot$ s (frequently called the "Lawson criterion"). Thus, if confinement is achieved for a few tenths to several seconds (typical of magnetic confinement), the plasma density required lies in the range of 10^{14} to 10^{16} ions/cm^3. If the confinement time is of the order of a thousandth of a microsecond (typical of inertial systems), much higher densities, of the order of 10^{26} ions/cm^3, must be achieved. In all cases, the plasma temperatures must be in the range of 10 to 15 keV.

Power Balance. To obtain useful amounts of power from fusion, not only must the output fusion power meet the conditions for power breakeven and all the other power inputs required by the reactor, but the fusion-released power must exceed this input power by a large amount, the difference being equal to the useful output. The power input to the auxiliaries of the fusion reactor, which heat and confine the plasma, is typically 10%–20% of the total power.

Since useful power output from a fusion reactor can be obtained only when substantial input power is provided, the reactor must be viewed as a power-amplifying device, the degree of power amplification differing according to the plasma-confinement scheme and the design of the lithium blanket, but typically is on the order of a few hundred.

5.4.4 Nonelectrical Applications

In addition to electrical production, fusion reactors can potentially be used for a variety of other important purposes. For example, neutrons from D-T (and other D-based fuels) can be employed very effectively in a fusion-fission hybrid to breed fissile fuel in the blanket region of the fusion device. A unique feature of these hybrids is that very high "support ratios" are conceivable; for example, with some concepts, as much as 20 times the energy production of the fusion plant can be provided in light-water fission plants. Fusion neutrons may also be used in conjunction with fission processes in hybrid systems and to accelerate the decay of fission wastes, that is, providing a fission waste burner. Several methods for chemical production by fusion plants have also been studied. These generally use the penetrating power of the high-energy neutron to create a high-temperature region in the blanket. This then appears well suited to various processes, for example, hydrogen production from water via either high-temperature electrolysis or thermochemical processes. The use of radiation, neutrons, or the plasma itself has also been considered for chemical processing, but this approach is not as straightforward (or as developed) as thermal processing.

Fusion driven propulsion units for deep space missions have also received serious attention. In this case, the fusion heated plasma can be exhausted through a magnetic nozzle to achieve directed thrust. This gives very high exhaust velocities (or high "specific impulses") making fusion propulsion extremely well suited for space missions to Mars and beyond where fast transit times are essential.

5.4.5 Plasma Confinement

The high temperatures and pressures of the fusion reaction preclude the use of a material vessel to confine the plasma during the reaction. Two primary approaches are now under investigation: magnetic and inertial confinement.

In the magnetic confinement scheme, intense magnetic fields are generated in the reactor, oriented with respect to the reaction space such that the plasma ions and electrons experience an inward pressure that resists the outward pressure of the hot plasma.

In inertial confinement, D-T-fueled targets (e.g., small plastic ampules) are bombarded by intense pulsed laser, x-rays, or charged-particle beams. Many beams are arranged to impinge symmetrically on the target from different directions; that is, beams are aimed at a common point at which the target is positioned. Each beam is pulsed and rapidly heats the surface of the target. A very high energy, exceeding a megajoule, is required in each pulse. Typically, each pulse lasts about 10^{-9} s and the pulse rate is about 1/s. The symmetrical heating at the surface of the pellet causes it to ablate. This creates a strong inward force such that the target's interior is compressed. Fusion then occurs at the center and the reaction burns outward. High-energy, heavy ions are considered to be potentially attractive for inertial confinement fusion, the key advantages being the high efficiency of these ion accelerators and improved coupling of the beam energy into the target compared to photons (lasers). Also, the potential for the scale-up of present accelerators to achieve economical, efficient, high-repetition-rate units seems good based on developing accelerator technology.

Magnetic Confinement. The principal means for magnetic confinement is the tokamak. Other approaches include, the reversed field pinch, the spheromak, the tandem mirror, and the stellarator. However, the level of support for research on alternative approaches is only a small fraction of that in tokamaks.

A measure for the performance of the magnetic confinement methods can be expressed by a parameter β, the ratio of the outward plasma kinetic pressure to the inward confining pressure of the magnetic field. This parameter denotes the efficiency of the magnetic confinement; that is, high-β systems make better use of the confining field than do the low-β systems. The parameter β is defined as

$$\beta = \text{constant} \times \frac{n(T_i + T_e)}{B^2}$$

where n is the plasma density, T_i and T_e the ion and electron kinetic temperatures, and B the confining magnetic field strength. High-β reactors are expected to operate with β greater than 0.8, and low-β designs at values below 0.2. The fusion power density varies as the square of β. Tokamaks are generally low-β devices. Spheromaks and reversed-field configurations are examples of high-β devices.

Inertial Confinement. Originally termed "laser fusion" this confinement method focuses a tremendous pulse of energy onto the surface of a sub- millimeter scale target containing fusionable fuel, e.g. D and T. This energy causes the outer layer of the target to be ablated away in "rocket exhaust" fashion. The resulting inward directed force causes compression of the target to ultra- high densities, approaching a thousand times solid density. In the process, compression heating raises the target fuel to fusion reaction conditions. A point is reached where the internal pressure of the hot, high density target just balances the internal compressive force. This results in a "stagnation" point in the inward trajectory of the target radius, followed by a rapid outward expansion. The time associated with this motion around a stagnation point can be viewed as the confinement time, τ. In effect the target dynamics is governed by the inertia of the individual target ions, hence the terminology "inertial confinement fusion". A rough estimate of τ is given by R/υ_{ith} where R is the compressed target radius and υ_{ith} is the average speed of the target ions. Then if R~10^{-6} m and υ_{ith}~10^3 m/sec, τ~10^{-9} sec. While very short, this confinement time proves to be adequate for net energy production.

The key point is that with the very high density, the fusion reaction rate is so large that this time is sufficient to produce more fusion energy than that required for the compression. In addition to lasers, it is now recognized that other pulsed energy sources (termed "drivers") such as x-ray producing Z-pinches and heavy ion accelerators can focus the required energy on these extremely small inertial confinement targets.

Other Confinement Methods. While magnetic and inertial confinement are the most widely considered approaches, a variety of other techniques have (or are being) considered. Examples include high speed acceleration of a target against a "wall" or another projectile, creating "impact" fusion. Electrostatic fields can not confine a plasma in steady state, but if directed ion motion is created, the ion inertia itself prevents directed plasma losses, providing Inertial Electrostatic Confinement (IEC). The replacement of an electron in liquid deuterium or tritium with a muon (i.e. a "heavy electron") results in a much smaller orbital radius for the muionic deuterium atom. This in tern allows atoms to approach each other at shorter distance than normal and thus permits fusion in the liquid (or in effect "cold fusion") without requiring the high ion velocities of a hot fusion. This is termed muon catalyzed fusion. These examples are far from being inclusive but they indicate the "richness" of the field.

5.4.6 Tokamaks

In tokamaks, the plasma-confinement space has closed, toroidal geometry, as shown in Fig. 5-24. A current in the toroidal field coils produces a toroidal field of the order of 100 kG, which confines the

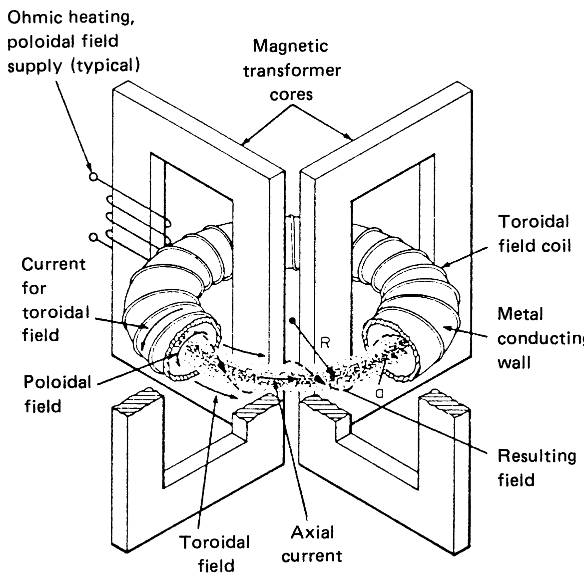


FIGURE 5-24 Elements of the tokamak reactor. Two magnetic fields (toroidal and poloidal) combine to produce a resulting field shown in the dashed line. This field both confines the plasma and causes its ions and electrons to move through the cavity. The ohmic heating of this axial current is one component of the heat required to achieve fusion of the plasma. Additional heating must also be supplied from an external source to maintain the fusion temperature in excess of 100 million degrees Kelvin.

plasma within the toroidal cavity. A second magnetic field, the poloidal field, is produced by external magnetic cores so that this field falls at right angles to the toroidal field everywhere within the cavity. The vector sum of the two fields (dashed line in Fig. 5-24) causes the plasma ions and electrons not only to be confined but to move through the torus, producing an axial current of tens of hundreds megaamperes. The ohmic heating associated with this current is an important source of power input for heating the plasma. However, in a tokamak, ohmic heating alone is insufficient to produce fusion temperatures. Additional plasma heating must be supplied by one of various methods, or combinations of them, such as radio-frequency (rf) heating, adiabatic compression, or the injection of high-energy beams of fuel particles.

A third, transverse, magnetic field is also needed to provide control of the position of the plasma column within the cavity. Additional magnetic coils may be involved to create a magnetic “divertor” which is used to scrape off the outer surface of the plasma and divert it into a “dump” vacuum pump region. The divertor serves three key functions: removal of impurity ions attempting to enter the plasma from the plasma chamber’s wall, removal of fuel and fusion products that diffuse to the outer surface of the plasma, and removal of heat carried by escaping plasma.

When operating in the burn cycle, the kinetic temperature of the plasma is expected to range from 10 to 30 keV; the latter figure is above 300 million degrees absolute. Several large experimental devices, for example, the JET tokamak in England and the TFTR tokamak at Princeton University, have achieved temperatures in the range of 10 keV, and near energy breakeven conditions. In pulsed operation of the tokamak, successive periods of plasma confinement will last several seconds or tens of seconds. The total burn time, prior to purging and reloading the reactor, will be of the order of several hundreds or thousands of seconds. Auxiliary heating power of 10 to 100 MW is required to produce a fusion power output of 1 to 5 GW. Recent studies have concentrated on the possibility of using rf or other techniques to provide a current drive which can ultimately permit steady-state rather than pulsed operation. Deuterium and tritium fuel would be injected at a rate of about 2 to 4×10^{22} atoms/s to provide a density of about 10^{14} ions/cm^3 in a low-β (3% to 10%) tokamak.

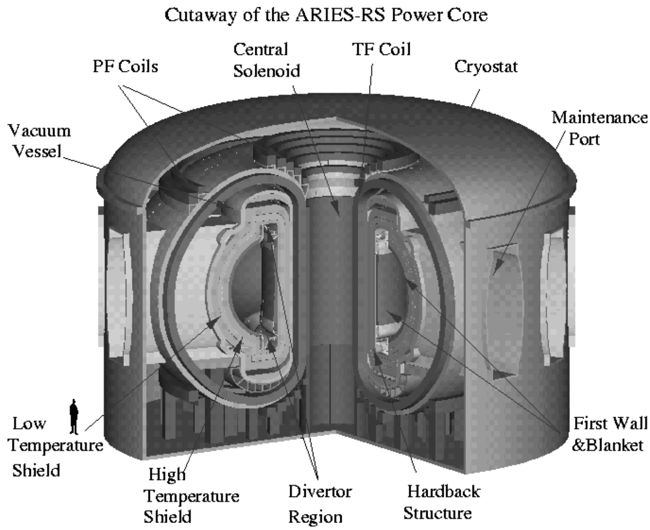


FIGURE 5-25 Schematic diagram of the ARIES-RS fusion power core unit. (*Najmabadi and the ARIES Team 1997.*)

A number of conceptual studies of the prospective operating characteristics of tokamak reactors have been made. Figure 5-25 and Table 5-6 are the results of such a study done by the ARIES study team at the University of California, San Diego, along with other university and industrial collaborators (Najmabadi 1997).

Figure 5-26 shows a schematic drawing of the TFTR which was designed to study burning D-T plasmas. Built at the Princeton Plasma Physics Laboratory (PPPL) at a cost of over $500 million, this

TABLE 5-6 Operating Parameters of the ARIES-RS Tokamak Power Plant

Aspect ratio	4.00
Major radius, m	5.52
Minor plasma radius, m	1.38
Plasma vertical elongation (*X* point)	1.70
Plasma current, MA	11.32
Bootstrap current fraction	0.88
Current-drive power, MW	81
Toroidal field on axis, T	7.98
Peak field at TF coils, T	16
Toroidal beta	0.05
Average neutron wall load, MW/m^2	3.96
Primary coolant and breeder	Natural lithium
Structural materials	Vanadium and steel
Coolant inlet temperature, °C	330
Coolant outlet temperature, °C	610
Fusion power, MW	2170
Total thermal power, MW	2620
Net electric power, MW	1000
Gross thermal conversion efficiency	0.46
Net plant efficiency	0.38
Recirculating power fraction	0.17
Mass power density, kWe/metric ton	66.70
Cost of electricity, mill/kWh	75.79

Source: Najmabadi and the ARIES Team (1997).

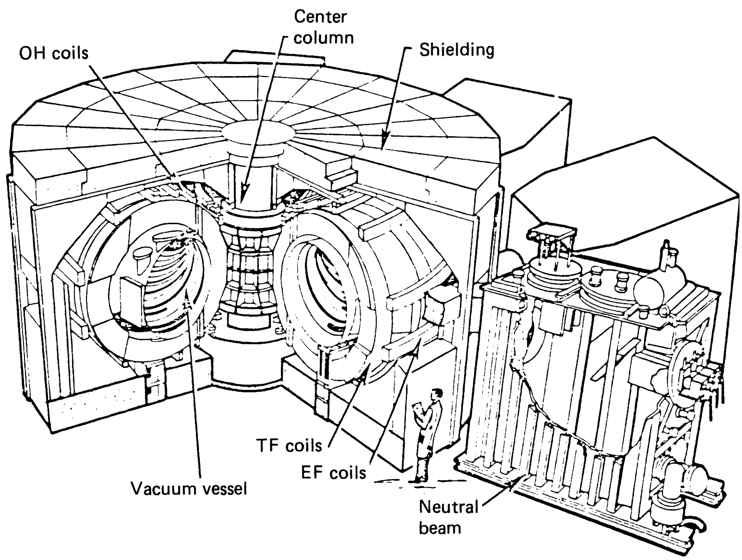


FIGURE 5-26 Schematic drawing of TFTR.

device began initial operation in December 1982. Since then, the TFTR has carried out a series of impressive experiments. Objectives have been to obtain improved confinement, a basic understanding and control of instabilities, improved heating and current drive, impurity control, and plasma-wall interactions. These results have been complemented by a number of other major tokamak experiments worldwide. These included D-III (San Diego, USA), JET (Culham, UK), JT-60 (Japan), and T-12 (Russia). A major step in this work involved the study of burning D-T plasmas in both TFTR and JET. This provided valuable experience with a fusion plasma where a major fraction of its heating comes from energetic fusion products (the 3.5-MeV alpha particle in the case of D-T fusion). It also provided experience with safe handling and control of tritium. Then, after several successful campaigns with D-T burns, Congress cut funding for TFTR stating it had met the design objectives. The experiment was then shut down in 1997.

The TFTR employed ~30 MW of neutral beam heating in a device with 12.48-m major radius (R) and a 0.85-m minor radius (a) and an on-axis toroidal field of 5.2 T. It produced plasmas with densities of 10^{13} particles/cm^3, plasma temperatures of 10 keV, and fusion power densities of the order of 1 MW/m^3. The JET experiment is still in operation.

5.4.7 World Facilities for Fusion Research and Reactor Concepts

A large number of both magnetic and inertial fusion research facilities are in operation worldwide. A complete list of institutes, ordered by country, city, and acronym, along with a listing of major equipment and scientific staff, is available from the International Atomic Energy Agency, Vienna (World Survey of Activities in Controlled Fusion Research, 2001 ed.).

Magnetic Confinement Facilities and Concepts

Operational Experience with Large Tokamaks. Extensive operational experiences have been obtained with hydrogen, deuterium, and D-T plasmas in major tokamaks around the world. Two—TFTR and JET—have used D-T plasmas. This experience, plus some data about prior operations, is provided here for these two major tokamak facilities.

TFTR operational experience is described in detail by von Halle (1998) and Johnson (1995).

In April 1997, the TFTR completed its final operating period, bringing to a close a highly successful phase of fusion research. The TFTR produced over 80,000 high-power plasmas since 1982 with the objectives of studying the plasma physics of large tokamaks, gaining experience in the solution of engineering problems associated with large fusion systems, and demonstrating fustian energy production from the burning, on a pulsed basis, of deuterium and tritium in a magnetically confined toroidal plasma system.

In 1993, TFTR became the first magnetic fusion device to study plasmas using nearly equal concentrations of deuterium and tritium. Since that time, over 1000 D-T experimental shots and over 23,000 D-D shots were carried out, demonstrating new regimes of plasma confinement, proof of alpha heating, and reactor-level fusion power densities by producing a plasma that yielded over 10 MW of fusion power at a corresponding central fusion power density of approximately 2.8 MWm^{-3} (Nazikian 1996, Bell 1997).

The TFTR technical systems routinely operated at or beyond the original design criteria throughout the period, maintaining an impressive machine availability of $>85\%$. This success continued through the final night of operations when a plasma with a record 7.7 MJ of stored energy was attained.

An important additional contribution from this work was the demonstration of safe operation of the tritium systems, with over 950 kCi of tritium processed within the constraints of a 50-kCi site limit and a 20-kCi machine limit. The tritium purification system "closed the loop" on the TFTR fuel cycle during the final operating campaign, processing >50 kCi of tritium at $>95\%$ purity back to the tritium storage and delivery system U-beds for further use on TFTR experiments.

JET operational experience is outlined by Bertolini and The JET Team (1997). JET started operation in June 1983, as the largest tokamak experiment of the coordinated fusion research program of the European Union. The global objective of JET was to produce and study plasmas of thermonuclear grade, in configurations suitable for extrapolation to a reactor. This led to the early choice of machine parameters with D-shaped toroidal coils, vacuum vessel, and plasma cross section. Great flexibility and suitable stress margins were included in the original JET design, to allow modifications and/or upgrading of the machine to follow the rapidly evolving requirements of the physics program (Keilhacker and The JET Team 1997).

Two major interventions took place between 1986 and 1989. The first involved increasing the plasma current capability from the design value of 4.8 to 7.0 MA in limiter configuration and from 3.0 to 5.0 MA in the *X*-point configuration. This work involved revising the design of the toroidal and poloidal coils, of the mechanical structure and of the vacuum vessel, by detailed finite element computer modeling and calculations, and fatigue tests on the prototype toroidal coil. Moreover, new power supplies had to be provided and plasma control also had to be improved, to cope with the enhanced vertical instability of the plasma ring. This upgrading was generated by initial experimental results, which showed a sharp decay of the energy-confinement time with heating. This could be counteracted by increasing the plasma current and by setting up an *X*-point configuration, which allowed H modes to be established.

The second change required progressively covering the inconel vessel walls with low-*Z* materials, specifically, graphite tiles ($Z = 6$) and later beryllium tiles ($Z = 4$), supplemented at first by wall carbonization and later by beryllium evaporation. This intervention was prompted by the need to substantially reduce Z_{eff} in order to decrease radiation losses.

At this point, the JET plasma could be maintained for only about 1 s, limited by a combination of MHD instabilities and accumulation of impurities in the *X*-point region. Active control of the impurities was therefore required, and was achieved by installing a pumped divertor to control impurity level and particle and energy exhaust and to enhance plasma energy confinement in H mode (Bertolini 1995). This represents the third major upgrading of the JET tokamak. The original features of the JET design allowed the basic structure of the machine to be maintained (viz., toroidal and poloidal coils, mechanical structure, and vacuum vessel). However, the necessity to install the divertor coils inside the vacuum vessel produced a loss of about 25% of the plasma volume, but it allowed plasma currents up to 6 MA to be accommodated.

Three divertor configurations (Mark I, Mark II, and Mark III Gas Box) have been designed, with progressively more closed configuration to enhance particle and impurity retention in the divertor chamber and to increase the amount of plasma energy released by radiation. The results of JET divertor studies are of great importance to finalize the ITER divertor design.

The full implementation of these measures led to the development of the hot-ion regime, showing a spectacular increase in overall plasma performance. The fusion triple product improved from

0.12 to 0.9 × 10^{21} m^{-3} (s · keV) and the equivalent energy gain Q_{DT} increased from 0.01 to 1.07. In separate pulses, an ion density of $n_D \sim 4 \times 10^{20}$ m^{-3}, an ion temperature T_D ~30 keV, and an energy-confinement time of $\tau_E \sim 1.8$ s, were obtained.

Finally, a $Z_{eff} \leq 2$ and a dilution factor $n_D/n_e \geq j0.9$ were reached. This level of global plasma parameters was sufficient to perform the first ever D-T experiment, which, in spite of using a mixture far from optimum (11% T–89% D), achieved a peak fusion power of 1.7 MW with >50% of thermalized neutrons, and a fusion energy of 2 MJ (The JET Team 1992). The additional heating power has been progressively increased up to 50 MW (20 MW of neutral beam injection, 20 MW on ion-cyclotron resonance heating, and 10 MW of lower hybrid). These results of internal plasma-wall components, especially the divertor, are under way with a new D-T experimental campaign planned for fall 1999 (The JET Team 1995).

A recent series of experiments with deuterium-tritium plasmas (DTE1) are of particular importance for fusion reactor developments (Stork and The JET Team 1997). These are the first plasma experiments with a 50:50 D-T mixture in a tokamak with a divertor.

Extensive work was undertaken to ensure that the JET machine and its major subsystems were able to safely carry out an extended period of D-T operation (Hemmerich 1989). This included adding the JET active-gas-handling system (AGHS) and its commissioning to full closed-cycle operation. As of 1998, the AGHS supplied around 40 g of tritium to the JET torus and neutral-beam injectors (NBIs) and reprocessed batches of tritium of over 11 g returned from the cryopumps of the torus and NBI.

Several modifications were required to bring the NBI system to full tritium compatibility (Stork and the JET Team 1997). The injection of tritium beams at energy up to 155 kV and total power of up to 11.3 MW has taken place.

Prior to D-T operation, JET was subjected to extensive deterministic analysis of design-basis accidents (DBAs) to establish changes required to protection systems and to satisfy the regulatory authorities.

A planned intervention to repair a small water leak in the tritium neutral injector was successfully undertaken during these operations (Stork and the Jet Team 1997). The exhaust detritiation system (EDS) was effective in keeping environmental discharges to below management limits throughout operations.

Other Major Tokamak Experiments. Although not all-inclusive, this section indicates the status and diversity of tokamak research worldwide.

The JT-60 tokamak, located at the Naka Fusion Research Establishment, Ibaraki-ken, Japan, is one of the world's largest tokamak experiments (Neyatani and The JT-60 Team 1995, Sakasai and The JT-60 Team 1997). To extend the earlier results from JT-60U, a superupgrade design (JT-60SU) has been developed. It has a superconducting toroidal field and poloidal field coil system with a pulse length of 2000 s or more. Figure 5-27 compares cross-sectional views of JT-60SU and JT-60U.

Typical operation parameters are shown in Table 5-7. The physics goal is a simultaneous achievement of stable steady-state full current drive plasma with high confinement and high bootstrap current fraction, together with a dense, cold, radiation-divertor function in reactor-relevant conditions.

For steady-state physics, the pulse length is set to exceed than 2000 s, which is sufficiently longer than the characteristic times of the particle saturation of the wall (30 to 60 s), current diffusion (100 to 400 s), and first-wall temperature (500 to 1000 s). Establishment of integrated physics, technology, and engineering for long-pulse operation under steady-state reactor-relevant conditions is also important for future steady-state reactors, such as developing superconducting coils, and high-heat-conductive secondary gamma-ray shielding. Progress in high-performance and steady-state experiments on JT-60U during 1995 to 1997 is summarized as follows:

> JT-60U research on the steady state with high fusion performance focused on contributions to ITER physics R&D. Experiments on negative-ion-based neutral beam injection (N-NBI) indicated a high neutralization efficiency of 60% at 370 keV and a high current drive efficiency of $\eta_{CD} = 8 \times 10^{18}$ $m^{-2} \cdot$ A/W.
>
> Toroidicity-induced Alfvén eigenmodes (TAE modes) excited by N-NBI were observed for the first time. The divertor was modified from the open divertor to the W-shaped pumped divertor in February to May 1997.

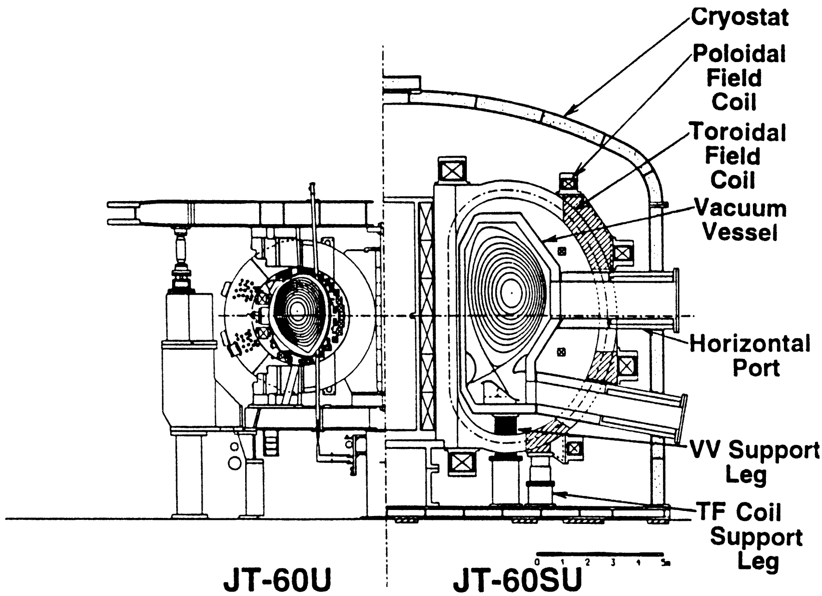


FIGURE 5-27 Cross-sectional views of JT-60SU and JT-60U.

In the new pumped divertor, radiative divertor by neon and deuterium injection with high recycling enhanced the divertor radiation with helium pumping in ELMy H-mode discharges.

Halo current was measured in the new divertor, and a new database was added for ITER. A high equivalent Q_{DT} of 1.05 was achieved in high-performance reversed magnetic-shear discharge.

Quasi-steady reversed-shear plasma with internal transport barrier and H-mode edge was sustained for 4.3 s. High-triangularity shaping was effective for improvement of the giant ELM limit and β_N. Steady-state high performance in the ELMy H mode was sustained for 9 s with the new divertor.

TABLE 5-7 Main Parameters of JT-60SU

Toroidal field strength B_T at 4.8 m, T	6.3
Plasma current I_p, MA	10
Major radius R_p, m	4.8
Minor radius a_p, m	1.3–1.5
Aspect ratio	3.2–4
Elongation κ	1.8
Triangularity δ	0.4–0.8
Number of TF coils	18
Total volt-seconds, V·s	170
Pulse length, s	>2000
NB power P_{NB}, MW	65
LHRF power, MW	20
ECRF for breakdown, MW	1

The DIII-D Tokamak. The mission of the DIII-D national program, centered at General Atomics in San Diego, Calif., is to establish the scientific basis for the optimization of the tokamak (Calis et al. 1989, Petersen and The DIII-D Team 1997). The research is carried out in four areas—transport, stability, boundary, and current drive and heating—in collaboration with a large number of national and international collaborators. The main goal is to optimize the performance of the tokamak through active control of the plasma shape and the plasma profiles. In the transport area, the emphasis has been on the study of the role of internal transport barriers in improving confinement. In the stability area, the emphasis has been on the study of neoclassical MHD, resistive wall stabilization, density limits, and disruption characterization and mitigation. In the boundary area, the emphasis has been on developing understanding of radiative divertor physics and the initiation of the new radiative

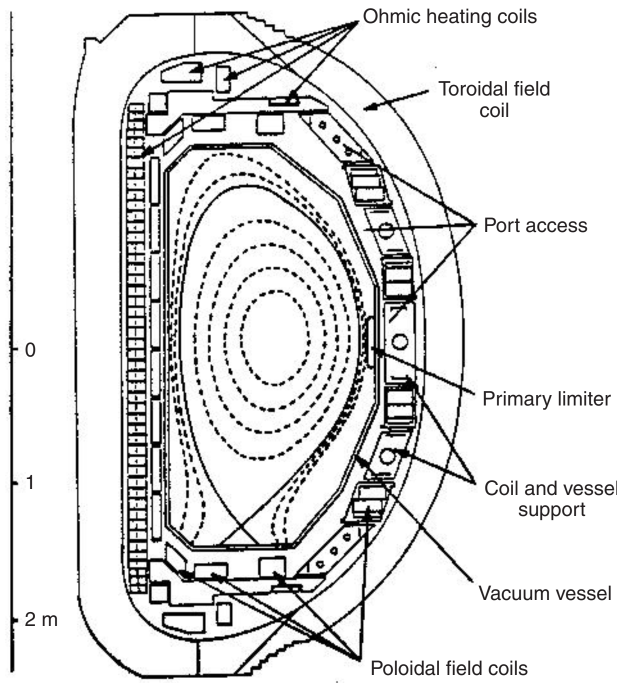


FIGURE 5-28 Cross section of the DIII-D device. Superimposed is the MHD equilibrium of a high-beta divertor.

divertor, and in the current drive and heating area on the preparation of two 110 GHz gyrotrons. These areas will be discussed in more detail in the following sections.

This research program utilizes the DIII-D tokamark (Fig. 5-28), which is a midsized device operating at near-reactor-level plasma parameters. It has a D-shaped cross section with an aspect ratio of 2.5:1 and a major radius of 1.6 m, and is capable of producing a large variety of plasma shapes, including elongation up to 2.6 and triangularity to 1.0. The maximum toroidal field on axis is 2.1 T, and the maximum plasma current is 3 MA. The shaping flexibility is due to the 18 fielding shaping coils that are located close to the plasma and individually controlled. A divertor baffle and a cryopump are installed above the floor and at the ceiling of the DIII-D vessel to allow study of divertor physics and control the plasma density. The auxiliary heating of the plasma is provided by eight neutral beams producing 20 MW at 80 kV in deuterium, 1.5 MW (source) of 110 GHz electron cyclotron heating, and 6 MW (source) of fast-wave current drive.

More than 50 different diagnostics are used to probe the DIII-D plasma. The Thomson scattering system measures the electron temperature and density every 2 ms at 40 different locations including the divertor region. The rotational Stark effect diagnostic measures the local pitch angle of the magnetic field, and thereby of the plasma current density, and the radial electrical field. The charge exchange recombination diagnostic measures ion temperature, poloidal and toroidal rotation, impurity density, and radial electrical field. There are several diagnostics that measure fluctuations (beam emission spectroscopy, far infrared scattering, etc.). A digital control system is used to control the plasma shape, current, and density and the injection of auxiliary heating power.

Increasing electron cyclotron heating (ECH) has high priority. With the successful testing of the two gyrotrons and development of new diamond windows, it appears that long-pulse gyrotrons are now available. The ECH power will be used for current profile control, perturbation studies, and transport barrier control. An extension of the DIII-D pulse length can be obtained by installing a new return bus for the toroidal coil, with minor upgrades of the field-shaping coil supplies. A longer tokamak pulse duration is important for stability studies, extension of the advanced tokamak mode, and wall stabilization. Upgrade of the Thomson scattering system to include a central cord is important for transport and neutral density measurements, understanding of disruptions, and to aid active mode control of neoclassical MHD modes. The top divertor will be completed with inner and private flux baffles, which should help reduce core fueling of neutrals and thereby improve confinement and increase the electron temperature for a more efficient current drive. Later, the lower divertor will be upgraded for advanced tokamak operations.

Tore Supra. Among the large tokamaks in operation around the world, Tore Supra (Centre d' études de Cadarache, France) has the unique feature of a superconducting magnet which provides a permanent toroidal field. With its niobium-titanium coils operating at 1.8 K, the Tore Supra toroidal field is routinely set up for 4-T/(12-h/day) operation. Its main research directions are therefore concentrated on the control of multimegawatt plasmas during long duration, with the steady state as ultimate goal, ultimately involving steady-state operation of a large tokamak (R = 2.4 m, a = 0.8 m, I_p = 1.7 MA). Both lower hybrid and fast waves are used to drive the current. Enhanced performances related to current profile shaping have been intensively studied (Martin 1997). Consequently, the Tore Supra inner vessel is equipped with several water cooled elements which interact with the plasma to allow injection of several megawatts during several tens of seconds. A world record in injected energy was achieved in 1996 with 280 MJ during a 2-min shot. Very long evolution times were then put in evidence in plasma wall interaction physics.

On the basis of these results, an enhancement in the capability to handle large input powers and control the particles over long duration is now under way. This is the main motivation for the Composants Internes et Limiteur (CIEL) project, which consists mainly in an upgrading of the first wall components:

1. A new toroidal belt limiter in the lower part of the vessel, made from carbon fiber composite (CFC) brazed on copper tubes. This has been designed to remove continuously up to 15 MW of convective power.

2. A set of pumps to evacuate all types of gas species through the toroidal throat of this limiter, allowing improved density control.

3. A new water-cooled radiation screen to cover as much as possible of the inner vessel, to avoid uncooled parts interacting with the plasma.

The Alcator C-Mod Tokamak. The Alcator C-Mod, located at MIT, Cambridge, Mass., is a high-field, high-particle and high-power-density advanced divertor tokamak. It is one of the only five divertor experiments capable of plasma currents exceeding 1 MA. The primary research goals of the experimental program involve divertor research with reactor grade parameters, critical tests of both empirical and theory-based scaling laws for transport, and rf heating and current drive and their application to advanced tokamak research. Figure 5-29 shows a side view of the machine. Note that Alcator C-Mod is enclosed inside a cryostat so that the magnets can be cooled with liquid nitrogen (LN_2) to nearly 77 K. The liquid is circulated through the magnets and then drains back to a 300-ga sump, where it is recirculated.

The C-Mod vacuum vessel, unlike most other tokamaks, is used as a structural element designed to support the OH and EF coils. Since thick stainless steel construction is required, large currents can therefore flow. However, these currents have been successfully compensated for by the control system combined with an excellent set of magnet diagnostics. Table 5-8 lists some of the machine parameters currently obtained and also those expected over the next few years as upgrades are made to the first-wall hardware and rf systems are added.

KSTAR Tokamak. The KSTAR (Korea Superconducting Tokamak Advanced Research) project (Choi et al. 1979) is the major effort of the Korean National Fusion Program to design, construct, and operate a steady-state-capable superconducting tokamak. The project is led by Korea Basic Science Institute and shared by national laboratories, universities, and industry along with international collaboration. The key design features of KSTAR are major radius 1.8 m, minor radius 0.5 m, toroidal field 3.5 T, plasma current 2 MA, and flexible plasma shaping (elongation 2.0; triangularity 0.8; double-null poloidal divertor). Both the toroidal and the poloidal field magnets are superconducting

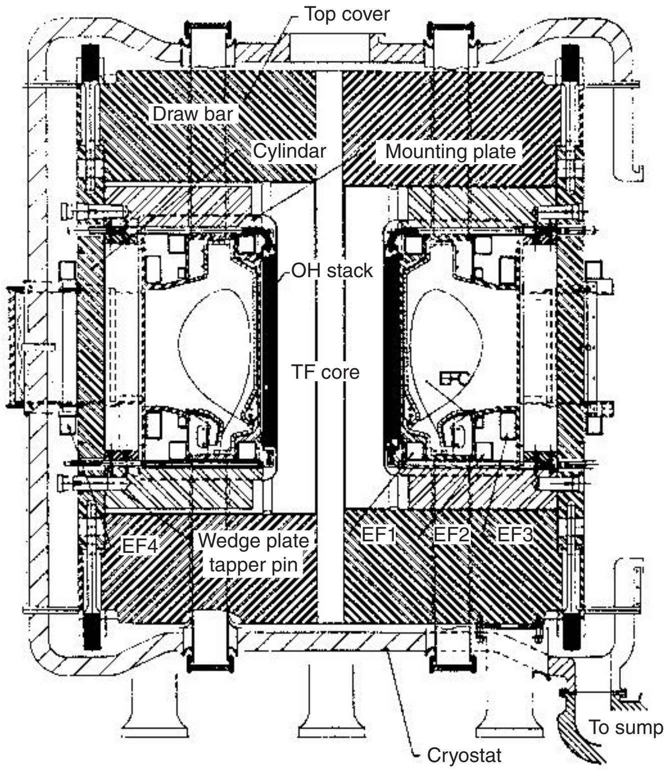


FIGURE 5-29 The Alcator C-Mod tokamak.

TABLE 5-8 Current and Future Alcator C-Mod Parameters

	Current	Future
Toroidal field, T	8	9
Plasma current, MA	1.5	2.5
Density, m^{-3}	$0.2–15 \times 10^{20}$	—
Elongation	1.8	—
Rf power, MW	4	12
Plasma volume, m^3	1	—

coils. The device is configured for upgrading for up to 300-s operation with noninductive current drive (Table 5-9).

The auxiliary heating and current drive systems consist of neutral beam, ICRF, lower hybrid, and ECRF. Deuterium operation is planned with a full radiation shielding. The KSTAR operational program is staged into two phases: the baseline and the upgrade. In phase I (2002 to 2005), inductively driven 20-s operation with the baseline auxiliary system (15 MW) is to be conducted, whereas 300-s operation with an upgraded auxiliary system (40 MW) for full noninductive current drive at high-beta regimes is planned for phase II (2006 to 2010).

The International Thermonuclear Experimental Reactor (ITER). The next major step expected in tokamak development has involved a massive international design effort. At the invitation of the director general of the IAEA, representatives of the world's four major fusion programs met in 1987 and developed a detailed proposal for a joint venture called the Conceptual Design Activity (CDA) for the proposed International Thermonuclear Experimental Reactor (ITER). The Director General then invited each interested party to cooperate in the CDA in accordance with the terms of reference developed at the meeting.

TABLE 5-9 Power and Current Drive Plus Projected Operation of KSTAR

Parameter	Baseline	Upgrade
Pulse length, s	20	300
Plasma heating power MW		
Neutral beam	8	24
Ion cyclotron	6	12
Lower hybrid	1.5	4.5
Electron cyclotron	0.5	TBD
Peak DD neutron source rate, s^{-1}	3.5×10^{16}	
Annual deuterium operating time, s	20,000	
Total number of pulses expected	50,000	

The ITER CDA, under the auspices of the IAEA, began in April 1988 and was successfully completed in December 1990. This work included two phases, the definition phase and the design phase. In 1988, the first phase produced a concept with a consistent set of technical characteristics and preliminary plans for coordinated R&D in support of ITER. The design phase produced a conceptual design, a description of site requirements, a preliminary construction schedule and cost estimate, and an ITER R&D plan. At the end of the design phase in 1998, however, global financial conditions forced a delay in the decision to proceed with construction of ITER. A reduced-size study produced a design for a smaller version at about half the cost.

The ITER design is based on a scientific knowledge base derived from the operation of dozens of tokamaks worldwide over the past decade. It also relies on the technical know-how flowing from the extensive fusion technology R&D programs of the four parties. The ITER tokamak is shown in Fig. 5-30. The principal parameters of the reduced-size version are listed in Table 5-10.

The main characteristics and parameters of ITER follow from its technical objectives which were derived to meet well-defined programmatic goals. Ignition requirements set the value of plasma current. Extended burn favors superconducting coil systems. The design goals for the first wall fluxes and fluence both dictate approximately the same minimum shield thickness. These requirements, combined with considerations of plasma stability, impurity control, and current drive, define the general features and approximate size of ITER. Nevertheless, within the freedom allowed by the technical objectives, the design philosophy has been to control size and minimize cost. Moreover, safety considerations are an integral part of the design activities.

The CDA defined both physics and technology R&D conducted in the various laboratories of the parties designed to validate the technical basis for the design. The physics R&D studies placed emphasis on reducing uncertainties in the divertor performance and energy confinement, extending the burn duration with noninductive current drive, and avoiding disruptions. The main categories of the technology R&D activities involved superconducting magnets, plasma-facing components, nuclear blankets and shields, remote maintenance, fueling systems, and heating and current drive systems.

The EDA phase lasted for 6 years and culminated in a well-documented basis for a decision on construction and siting. Construction would require about 8 years, followed by two "phases" of operation (physics and technology) totaling about 18 years, after which decommissioning would commence.

The EDA work involved about 1200 professional worker years with a team whose strength peaked at about 180 professionals. The cost of this design work, including full overhead and support personnel, was about $250 million.

The professional workerpower needed during the construction activities has been estimated to be about 2900 worker-years, primarily at the construction site. The actual construction cost of ITER has been estimated to be about $4.9 billion. During the construction phase, technology R&D would be needed (estimated to cost about $300 million). The cost of physics R&D directly attributable to ITER should be added to these costs, but this is difficult to separate out from generic fusion R&D in progress at the same time.

A central team of 300 professionals would be needed during the operation activities, including physicists and engineers, to operate the machine and supervise the testing program. The resulting

TABLE 5-10 Main Parameters and Dimensions of ITER

Total fusion power, MW	500–700
Neutron wall loading, MW/m^2	0.7
Plasma inductive burn time, s	300
Plasma major radius, m	6.2
Plasma minor radius, m	2.0
Plasma current I_p, MA	16
Vertical elongation at 95% flux surface	1.7
Triangularity at 95% flux surface	0.32
Safety factor at 95% flux surface	3
Toroidal field at 6.2 m radius, T	5.3
Auxiliary heating power, MW	73

Source: ITER (U.S.) Home Team Group. *ITER Final Design Report*, IC-33, Project Office, University of California, San Diego, 1998.

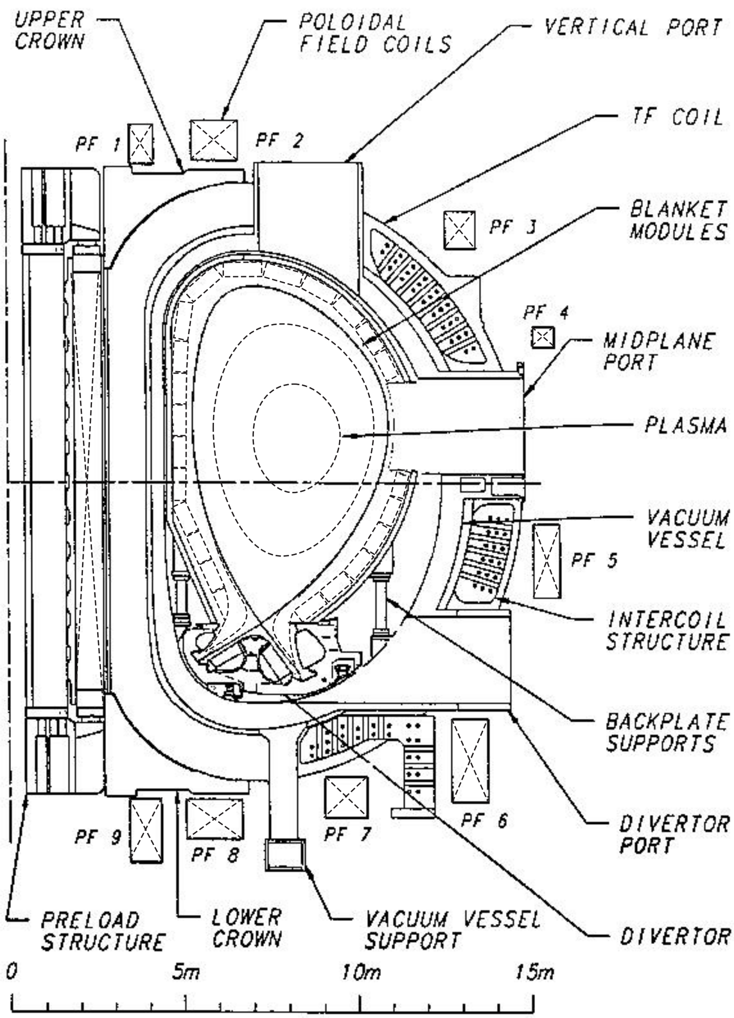


FIGURE 5-30 Cross section of the ITER tokamak.

preliminary operating cost estimates amount to about $270 million per year. The cost of the experimental part of the ITER program would be in addition to this basic operational cost. Decommissioning costs have not been estimated.

In conclusion, the ITER project has been a pioneering program in both technology and international collaboration. The goal for ITER has been to provide the basis for a demonstration reactor intended to open the way to commercialization of fusion power through the "tokamak route." Proponents of alternative magnetic confinement approaches and inertial confinement fusion (ICF) have been critical of the size and cost of ITER. They propose alternative approaches which they claim are more attractive and faster to develop. However, the extensive worldwide effort on tokamaks combined with the large experimental database accumulated from the many tokamak experiments carried out thus far have provided the tokamak with the lead in this development to date.

ITER construction is scheduled for 2006 with initial operation in 2014. Two sites have been proposed, one in France and the other in Japan.

Alternative Magnetic Fusion Concepts. Table 5-11 summarizes a representative cross section of alternative fusion concepts (AFCs) that, in one way or the other, have been or are being considered for the production of electrical power, chemical process heat, and/or fissile material. Depending on the confinement scheme considered, systems studies of AFCs range from a simple physics analysis, based on Lawson-like criteria, to detailed conceptual designs. With few exceptions, most reactor studies of alternative concepts fall into the less formalized part of this spectrum. For this reason, a quantitative intercomparison and ranking is not possible.

The toroidal AFCs summarized in Table 5-11 are classified as quasi-steady-state, long-pulsed (10 to 100 s), and pulsed (~1 s). A sampling from each category is given.

National Spherical Torus Experiment (NSTX). The National Spherical Torus Experiment (Fig. 5-31) is designed to prove the scientific principle of the spherical torus (ST) plasma. The ST plasma is nearly spherical in shape; its minor radium is slightly smaller than its major radius, thus giving an aspect ratio close to 1. ST plasmas may have several advantageous features such as a higher pressure for a given magnetic field. Since the fusion power density is proportional to the square of the plasma pressure, the ST is an example of an innovative alternative fusion concept that could lead to smaller and more economical sources of fusion energy.

Princeton Plasma Physics Laboratory (PPPL) leads the project and operates the NSTX facility, located at PPPL. The design and construction of the NSTX is a joint project of PPPL, the Oak Ridge National Laboratory (ORNL), Columbia University, and the University of Washington. The NSTX will have modular components for ease of repair and upgrade.

A national research team will carry out the research program on NSTX, which will cover a broad range of fusion and plasma science topics. The ST plasma possesses features of tokamak and spheromak plasmas, as seen in Fig. 5-32. The magnetic-field line of the ST plasma resembles that of the tokamak plasma at the stable, inboard side and that of the spheromak at the unstable, outboard side. This leads to a number of attractive physics features of interest to fusion energy science, estimated according to our present understanding in toroidal fusion science. These include

- MHD-stable plasmas with very high average toroidal beta (plasma pressure divided by externally applied toroidal field pressure at the major radius) in the range of 25% to 50%.
- Nearly complete (greater than or equal to 90%) alignment of self-driven current profile with the required plasma current.
- Strong magnetic shear and magnetic well (~30%), which help stabilize plasma microinstabilities believed responsible for plasma turbulence and rapid energy loss.
- Increased flow shear (by up to two orders of magnitude beyond tokamaks), which tends to suppress the remaining plasma microinstabilities.
- High plasma dielectric constant that permits a new class of potentially efficient rf-heating and current drive techniques, such as the high-harmonic fast wave and the electron cyclotron wave conversion to electron Bernstein wave.

TABLE 5-11 Summary of Alternative Concepts for Magnetic Fusion

Mirror		
	Simple mirror	
		Baseball coil
		Yin-yang coil
	Field-reversed mirror	
	Tandem mirror	
Quasi-toroidal advanced systems		
Advanced toroidal systems		
	Quasi-steady-state	
		Spherical tokamak
		Stellarator
		Helitron
		Torsatron
		Bumpy torus (EBT)
		Toroidal bicusp (Tormac)
		Surface magnetic confinement (Surmac)
	Long pulsed	
		Reversed-field pinch (RFP)
		Ohmically heated torus (OHTE)
		Ohmically heated tokamak (Riggatron)
	Pulsed	
		Theta-pinch (RTP)
		High-β stellartor (HBS)
		Belt-shaped screw pinch (BSP)
Compact toroid		
	Stationary	
		Spheromak
		Field-reversed mirror (FRM)
		Triggered-reconnected adiabatically compressed torus (TRACT)
		Electron-layer field-reversed mirror (Astron)
		Slowly imploding liner (LINUS)
	Translating	
		Spheromak
		Field-reversed theta pinch (CTOR)
		Moving-ring field-reversed mirror (MRFRM)
		Ion-ring compressor
Linear		
	Steady-state	
		Multiple-mirror solenoid
	Pulsed	
		Linear theta pinch (LTP)
		Laser-heated solenoid (LHS)
		Electron-beam heated solenoid (EBHS)
Very dense (fast-pulsed, linear) systems		
	Fast-imploding linear (FLR)	
	Dense plasma focus (DPF)	
	Wall-confined shock-heated reactor (SHR)	
	Dense Z-pinch (DZP)	
	Passive liners	

- Minimized plasma magnetic flux and helicity content to ease noninductive plasma start-up, including coaxial helicity injection or bootstrap current overdrive.
- Naturally diverted plasma scrape-off layer with large flux expansion (up to 10 or more) and magnetic mirror ratio (up to 4:1).

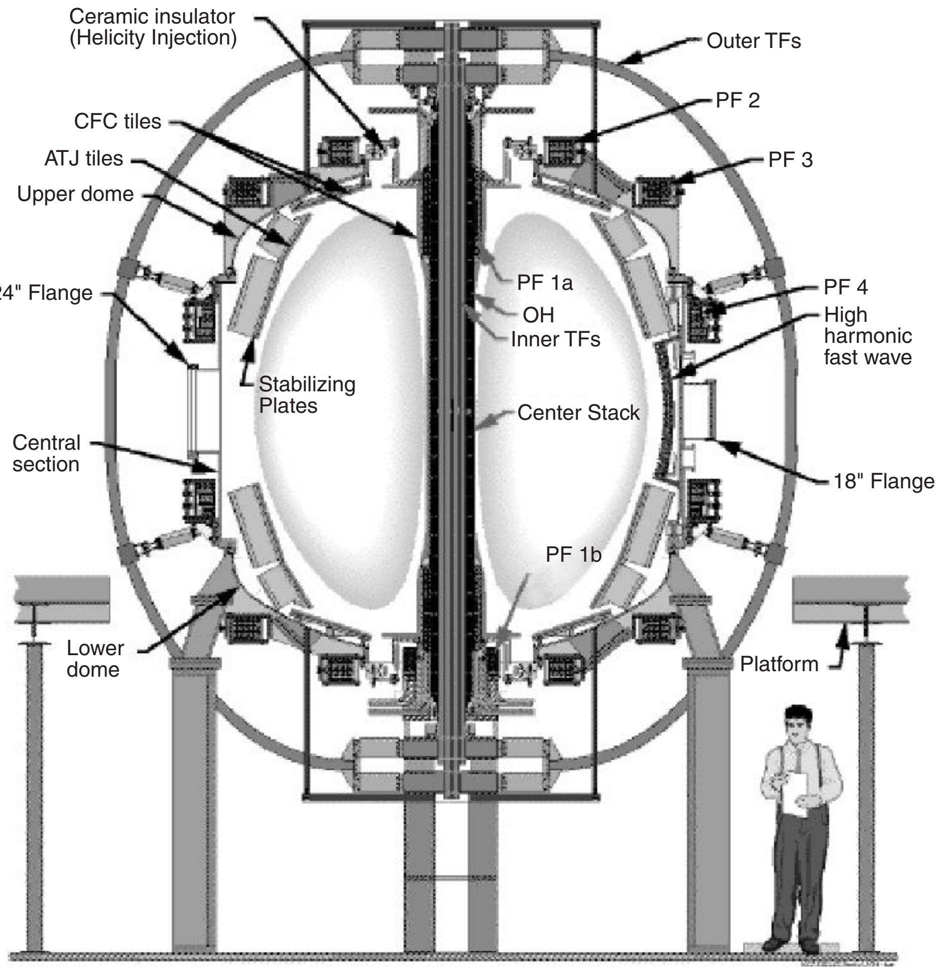


FIGURE 5-31 Cross section of the National Spherical Torus Experiment (NSTX).

These interesting features indicate great opportunities for discovery, innovation, and advancement in fusion and plasma sciences. If verified, the ST plasma will open up new opportunities for attractive applications of magnetic fusion using compact devices.

Figure 5-33 indicates how the ST plasma, given successes in proof-of-principle and proof-of-performance tests, could lend itself to a pathway for developing fusion power.

The NSTX effort grew from successful pioneering experiments conducted worldwide since 1991. These include the START (small tight-aspect-ratio tokamak) at the United Kingdom Atomic Energy Authority-Fusion, UK; the CDX-U (current drive experiment—upgrade) at PPPL; the HIT (helicity injected tokamak) at the University of Washington; the TS-3 (Tokyo Spheromak-3) at the University of Tokyo, Japan; and the SPHEX (spheromak experiment) at the University of Manchester Institute of Science and Technology, UK.

The data from these experiments have shown that the ST concept is ready for proof-of-principle tests. The NSTX is one of the new generation of experiments for this purpose. Other experiments include the MAST (megaampere spherical tokamak) at the United Kingdom Atomic Energy

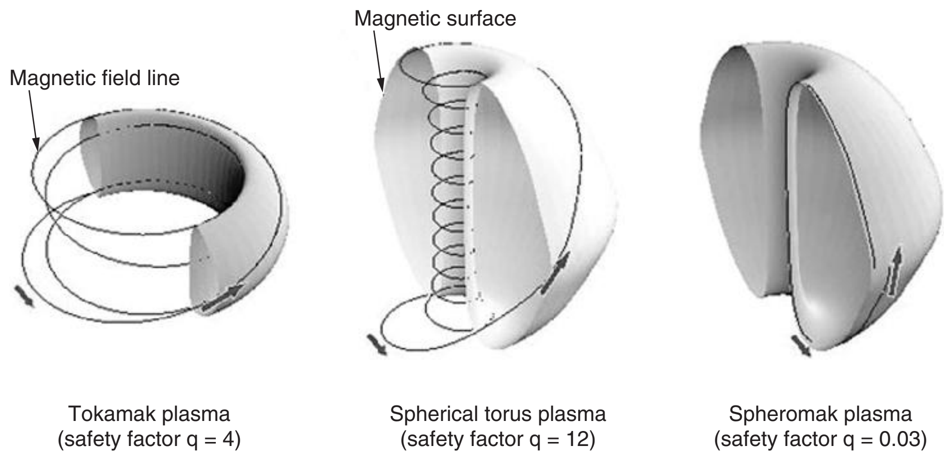


FIGURE 5-32 Spherical torus plasmas combine advantageous features of tokamak and spheromak plasmas.

Authority-Fusion, UK; the GLOBUS-Modified at the Ioffe Physical Technical Institute of St. Petersburg in the Russian Federation; and the Pegasus at the University of Wisconsin. These began experimentation during the second half of 1998.

These experiments will have strong complementary capabilities, and when combined, will permit investigations for plasma currents in the range of 1 MA, stability, toroidal β limits approaching

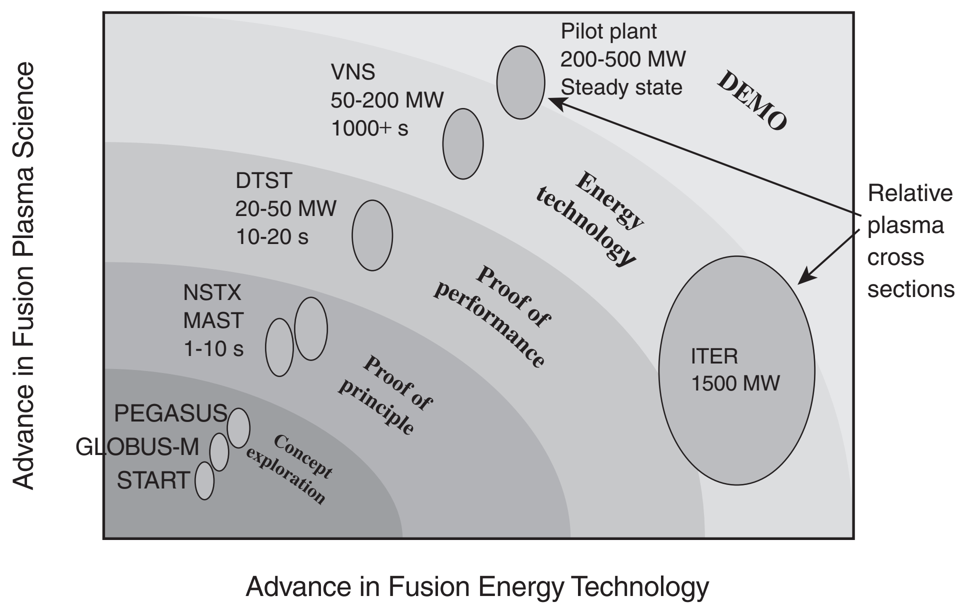


FIGURE 5-33 Schematic diagram of fusion science and technology advances.

50%, bootstrap currents approaching 90% of the plasma current, aspect ratios as small as 1.1, elongation as high as 4, and duration up to 5 s.

Plasma-heating and current-drive techniques to be tested cover neutral-beam injection, a full range of rf waves, and the more exotic noninductive start-up techniques.

The NSTX Research Program will investigate the fusion science principles of the ST plasma, covering

- Noninductive start-up, current sustainment, and profile control
- Confinement and transport
- Pressure limits and self-driven currents
- Scrape-off layer and divertor physics
- Stability and disruption resilience

These principles will be investigated in the scientifically interesting regimes that are relevant to near-term applications such as volume neutron sources (VNS) and fusion pilot and power plants of the future. Thus, the following regimes will be stressed:

- High average toroidal betas (25% to 45%)
- High pressure-gradient-driven current fractions (40% to 90%)
- Fully relaxed, noninductively sustained current profile
- Collisionless plasmas with high temperatures and densities
- Aspect ratios (major radius over minor radius) as small as 1.25 and elongation above 2

The NSTX device (Fig. 5-31) is designed to operate with the baseline parameters in Table 5-12 for plasmas with several forms of divertor and limiter configurations. The NSTX facility shall deliver to the plasma and handle the following power for start-up, heating, and current drive:

- 6 MW of high harmonic fast wave (HHFW) for up to 5 s
- 2 MJ of coaxial helicity injection (CHI) during 5 to 10 ms
- 20 kA of CHI injected current for up to 5 s
- 20 kW of electron-cyclotron heating (ECH) for up to 5 ms

The NSTX facility design will further account for the possibility of adding and handling 400 kW of ECH for up to 100 ms for noninductive start-up and 5 MW of neutral-beam injection (NBI) for up to 5 s for additional heating and current drive. The center stack of NSTX is highly compact to permit a plasma aspect ratio of as low as 1.25. It contains the inner legs of the toroidal field coils, a solenoid to provide full capability for inductive operation, a pair of poloidal field coils most effective in varying the plasma triangularity, the inboard vessel wall containing these components, the protective

TABLE 5-12 National Spherical Torus Experiment Baseline Parameters

Parameter	Value
Plasma major radius R_0, m	0.85
Plasma minor radius a, m	0.68
Applied torodial field B_t kG	3.0 at R_0
Plasma current I_p, MA	1.0
Plasma elongation κ	Greater than or equal to 2
Plasma triangularity δ	Approximately 0.4
Edge safety factor q	Approximately 10
Plasma pulse flat-top length τ_{pp}, s	About 0.5 (inductive-only operation) about 5.0 (driven-current operation)

graphite tiles, and the interface to the remainder of the device. The center stack is designed to be replaceable without affecting the rest of the device.

The vacuum vessel is 3.6 m in height and accommodates a plasma elongation of up to 2.2 and a triangularity of up to 0.6 for an aspect ratio of 1.25; a higher elongation of about 3 can be obtained for an aspect ratio of 1.4 by operating plasmas with a reduced minor radius.

Auxiliary heating and current drive for current profile control address another essential part of the NSTX mission. High harmonic fast-wave operation is expected to effectively accommodate the high-beta and high-bootstrap-current fraction plasma discharges and fully utilize rf power sources available on-site. A conducting shell, also protected by graphite tiles, will be placed near the outboard plasma edge to enable the investigation of maximized MHD stability limits in plasma beta and bootstrap current.

Mirrors. Mirror-type fusion reactors are characterized by open magnetic-field geometries; that is, magnetic flux passes through a mirror-type device and intersects material walls outside the reaction chamber. In order for such a device to be an adequate container of fusion plasma, it is essential that the end leakage of the plasma be strongly inhibited.

The earliest magnetic mirror configuration used a solenoid with increased magnetic-field strength near its end (Fig. 5-34). In this "simple" mirror, a charged particle travels in a helical orbit around an axially directed magnetic-field line. When traveling into a region of increasing magnetic-field strength, most particles are reflected, that is, their axial motion is stopped and reversed before they can penetrate to the high point of the mirror field. (Some particles with highly directed velocity may still escape. Strong magnets can reduce this leakage, but it cannot be completely eliminated.)

The plasma in a simple mirror is unstable to sideways motion because the magnetic field weakens in directions perpendicular to the coil axis. (The radial weakening of field is evidenced by the concave-inward surface of the magnetic flux bundle.) To solve this stability problem, the simple mirror configuration was replaced by the minimum-*B* mirror. From the center of this field—produced by a pair of

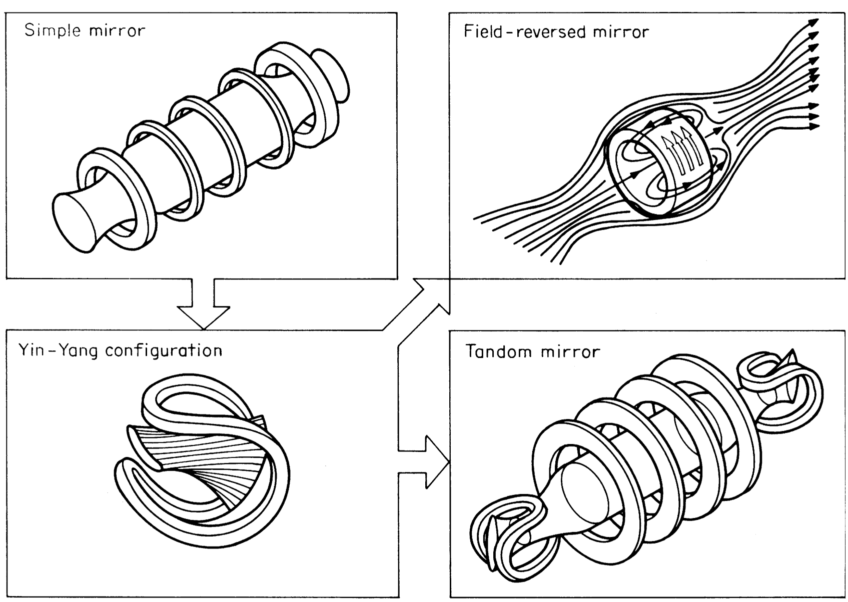


FIGURE 5-34 Evolution of mirror confinement concepts. The field-reversed mirror uses an internal ion ring (created by injection and by diamagnetic currents) to effectively create a torus configuration embedded inside the open mirror field. The tandem mirror uses electromagnetic fields created by ion injection to reduce the leakage from the ends of a solenoidal field.

solenoids and Ioffe bars, a baseball coil (shown in Fig. 5-34), or yin-yang coil—the field strength increases in all directions. However, by 1975, it was concluded from conceptual design studies of mirror reactors based on the minimum-B geometry that end losses from such a reactor would severely limit its plasma Q (fusion power divided by trapped injected power). The subsequent search for enhanced-Q mirror machines led to two new concepts: the field-reversed mirror and the tandem mirror.

In a field-reversed mirror (FRM), the confinement of plasma occurs through a toroidal region of closed magnetic-field lines generated by plasma currents in a nearly uniform background field (Fig. 5-34). The FRM offers the exciting possibility of fusion electric power reactors in small sizes. For example, conceptual design studies were carried out for both a multicell reactor producing 75 MW (electric) of net electric power and a single-cell pilot reactor producing ~11 MW (electric). However, following a series of initial experiments at the Lawrence Livermore National Laboratory (LLNL) that failed to build up sufficient plasma current for reversal, most of the LLNL effort shifted to the tandem mirror.

The tandem mirror confinement concept, invented in 1976, received strong study until cancellation of the U.S. program in 1985. The basic concept entails the improved axial confinement of a long cylindrical fusion plasma within a solenoid by means of strong electrostatic potentials at the ends, produced by mirror-confined, end-plug plasmas (Fig. 5-34). An improvement in the tandem mirror involved use of a *thermal barrier*. This concept uses enhancement of the confining electrostatic potential to establish a hotter electron population in the plugs than in the central cell. However, in the normal operation, electron flow between the plugs and central cell presents large temperature differences. Thus, a thermal barrier is necessary to reduce the passing of central-cell electrons into the plug. Basically, the thermal barrier consists of a region of much reduced magnetic field strength, plasma density, and plasma potential. This causes a depressed positive plasma potential which serves as an electrostatic barrier to electrons.

In order to maintain the thermal barrier, however, it is necessary to prevent the filling of the barrier region by thermal ions leaking from the central cell. Several methods of "barrier pumping" have been investigated. In one, the pumping is accomplished by trapped ions undergoing charge exchange interactions with axially directed neutral beams.

A preliminary conceptual design of a power reactor based on the tandem mirror has been carried out, including use of thermal barriers. The D-T fusion plasma is contained in the 56-m-long central cell and produces 1770 MW of fusion power. With $Q \simeq 10$, the reactor produces ~500 MW of net electricity. The central cell consists of twenty-eight 2-m-long modules, each containing an annular blanket region, a magnetic shield region, and two niobium-titanium solenoidal magnets. The entire central cell resides in a vacuum trench, which allows the module-to-module seal to be made by an annular metal inflatable cushion. The plug plasmas, contained in the plug yin-yang coils, are each sustained by a low-current, 400-keV neutral beam. A gyrotron tube system is used for microwave heating of the electrons on the plug side of the thermal barrier. Neutral beams on the end wall of the plug vacuum vessel provide charge-exchange pumping of the barrier and fueling of the central cell.

Advanced Toroidal Systems. Various "advanced" toroidal confinement systems are under active study. The spherical tokamak directed earlier represents one approach. Others include the stellarator, (and torsatron) and the helitron.

Stellarator-Torsatron. Unlike the tokamak, the nonaxisymmetric stellarator achieves equilibrium in a toroidal geometry by externally inducing a rotational transform in the confining magnetic field (Fig. 5-35); ideally, no axial currents need to be supported by the toroidal plasma column, as is required in a tokamak, although until very recently all stellarator experiments utilized such currents for ohmic heating. Implementation of a deformation (twist) into a simple toroidal field coil set allows the Torsatron magnetic geometry to be produced while eliminating the helical coil set in favor of a highly modular device (Fig. 5-36). In addition, more optimally oriented coil forces and lower stresses are anticipated for this modular Torsatron approach. These new advances have renewed interest in the reactor extrapolation of the stellarator-Torsatron concept, a renaissance that coincides with experimental success in heating a low-ohmic-current device, the latter being a prerequisite for eventual steady-state reactor operation.

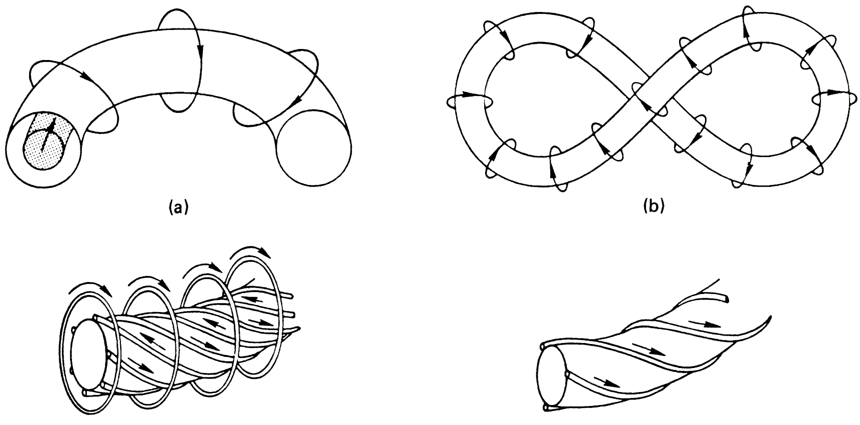


FIGURE 5-35 Toroidal configurations. (*a*) Tokamak: the main magnetic field is provided by poloidal coils with the poloidal field due mainly to a toroidal current in the plasma. Several external toroidal coils (which are not shown) provide a vertical magnetic field as well as drive the plasma current. (*b*) Figure-eight stellarator: only one set of solenoidal coils is needed. The rotational transform is generated by the torsion of the magnetic axis. (*c*) Classical stellarator: a set of 2*l* helical coils with current flowing in opposite directions inside the toroidal field coils provides the rotational transform. (*d*) Torsatron: a single set of *l* helical windings provides both the toroidal and poloidal fields. Usually, an additional set of toroidal coils (which are not shown) is necessary to provide an additional vertical field. The standard heliotron configuration is similar to this, but has an additional set of toroidal field coils.

Qualitative advantages that can be invoked for the Torsatron reactor concept include steady-state magnetic fields and burn, operation at ignition or high Q for low recirculating power, impurity and ash removal by means of a magnetic limiter and helical poloidal diverter that occur as a natural consequence of the topology, and no major plasma disruptions that could lead to an intense, local energy dump and no auxiliary position or field-shaping coils and moderate aspect ratio (>10), both of which ease maintenance access.

Stellarator Experiments. Next to tokamaks, the largest worldwide effort in magnetic fusion is focused on stellarator (Boozer et al. 1998, Wakatani 1998). The stellarator concept offers a potentially attractive alternative to the tokamak approach in magnetic confinement fusion. The magnetic field structure, which is required for equilibrium and stability of the confined plasma, is generated by external coil currents only and has an inherent steady-state capability. The rotational transform is generated in the classical approach by helical windings surrounding the plasma vessel. Plasma confinement under net current free operation has been demonstrated in experiments at Garshing, Germany. The net current free operation avoids disruptions.

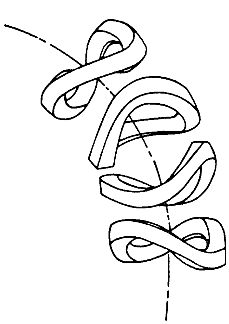


FIGURE 5-36 Coils in a typical modular stellarator.

Two small experiments are located in the United States; HSX at the University of Wisconsin and CAT at Auburn University (major radii of 1.2 and 0.5 m, respectively). Two new major stellarator experiments: LHD and W7-X will provide integrated tests of particular stellarator configurations using superconducting coils. Divertor, transport, and β-limit issues are being studied on CHS in Japan and W7-AS in Germany. TJ-II in Spain (1997) and H-1 in Australia focus on β-limit issues. Stellarator research is also being pursued in Russia and the Ukraine. The theory programs associated with the major stellarators are focused on support for the experiments. Longer-range projects include a better free-boundary package for the MHD stability codes, which is under development at Garching,

TABLE 5-13 Main Parameters of the Last Three Successive Stellarators at IPP

	R, m	a, m	A	B_0, T	Shear, %	Iota (i)	Coil type
W7-A	2	0.13	15	3.75	2	0–0.6	Helical + toroidal
W7-AS	2	0.18	11	3.0	±2	0.25–0.6	Modular + planar normal conducting
W7-X	5.5	0.55	10	3.0	15	0.8–1.2	Modular + planar superconducting

Germany. Studies are also starting on the implications of different stellarator configurations for a fusion power plant. A new stellarator, NCSX, in under construction at Princeton.

The Wendelstein 7-X (W7-X) Stellarator. The Wendelstein 7-X Stellarator (W7-X) is the next-step device in the stellarator line of the Institute for Plasma Physics (IPP) in Garching, Germany. A new branch of IPP is in Greifswald, Germany and will house the W7-X.

The stellarator line of IPP concentrates on low-shear devices, which exclude resonances with low rational numbers $n/m = {}^1/_3, {}^1/_2 \cdots$ from the confinement region and provide stability by a magnetic well rather than by strong shear. The classical stellarator approach such as W7-A, however, has enhanced neoclassical transport by trapped particles, which is not tolerable in reactor-scale devices. Furthermore, the β limit is too low for an economic power reactor, which requires a β value of typically 4% to 5%. The pressure-driven bootstrap current generates a significant contribution to the total rotational transform in classical stellarators, and thus contradicts the concept that the magnetic parameters can be controlled by external currents alone. These considerations have led to the next-step W7-X device. Some main parameters of the last three successive stellarators at IPP are shown in Table 5-13, where the stepwise development toward W7-X becomes obvious.

The coil and magnetic configuration for W7-X are depicted in Figure 5-37. Fifty nonplanar coils provide the confining magnetic configuration. The flux surfaces show a variation from a strong

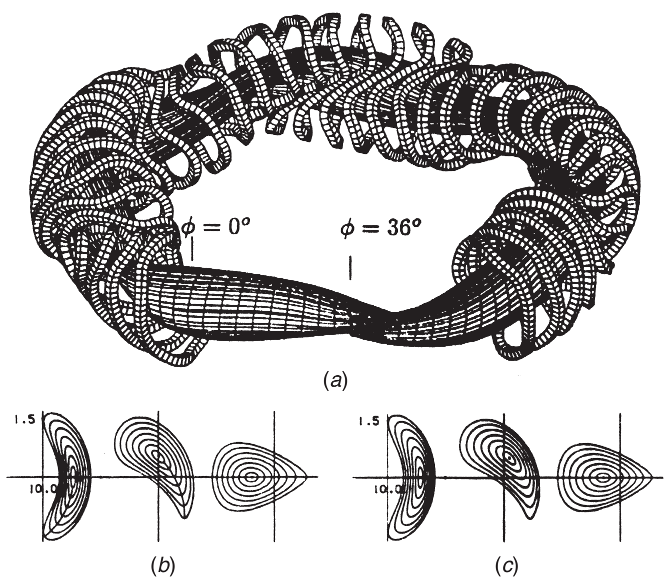


FIGURE 5-37 (*a*) The coils system of W7-X—50 nonplanar coils provide the confining magnetic configuration; (*b*) flux surfaces at three positions along one-half magnetic field period for $\beta = 0$; (*c*) $\beta = 5\%$.

indented elongated shape (bean shape) in the $\phi = 0°$ plane to a triangular shape in the $\phi = 36°$ plane. The standard configuration has a "built-in" profile of the rotational transform with $\tau = 0.86$ on axis and 1.0 at the edge. As is known from W7-AS experiments, the confinement depends strongly on the edge value of the rotational transform. For this reason, experimental flexibility is built into W7-X to allow the variation of the rotational transform and the profile in a wide range. A separate set of planar coils can be independently powered and allows a variation of the rotational transform from 0.75 to 1.01 on axis and 0.83 to 1.25 at the edge. Thus, both low- and high-shear configurations can be verified. The physics goals for W7-X include

1. Demonstration of quasi-steady-state operation in a reactor-relevant plasma parameter regime, with temperatures $T_e < 10$ keV, $T_I = 2$ to 5 keV, and densities $n_e = 0.1$ to 3×10^{20} m^{-3}
2. Demonstration of good plasma confinement to improve the database for reactor extrapolation
3. Demonstration of stable plasma equilibrium at a reactor-relevant plasma β of about 5%
4. Investigation and development of a divertor to control plasma density and impurities

W7-X does not aim at D-T operation, and provisions for remote handling in radioactive environments are not foreseen.

The Large Helical Device (LHD). The LHD at the National Institute for Fusion Science, Nagaya, Japan, is a superconducting (SC) toroidal fusion facility, is a heliotron-type device and has $l = 2$ SC continuous coils and three sets of SC poloidal coils. LHD has a maximum stored energy of 1.6 GJ at 4-T operation (Motojima 1993, 1995). The dumbbell-shaped vacuum chamber is installed in the inside of SC coils and a built-in helical divertor with the double-null structure functions for producing the steady-state plasma. The schematic view and basic parameters of LHD are shown in Fig. 5-38 and listed in Table 5-14. The volume and surface area of the vacuum chamber are 150 m^3 and 850 m^2, respectively.

To ascertain and realize the steady-state operation of LHD, several important issues concerned with both physics and technology should have been carefully considered and fixed during the design and construction phases. Of interest here are the major operations scenarios of steady-state plasmas based on both physics and technology aspects, especially divertor functions.

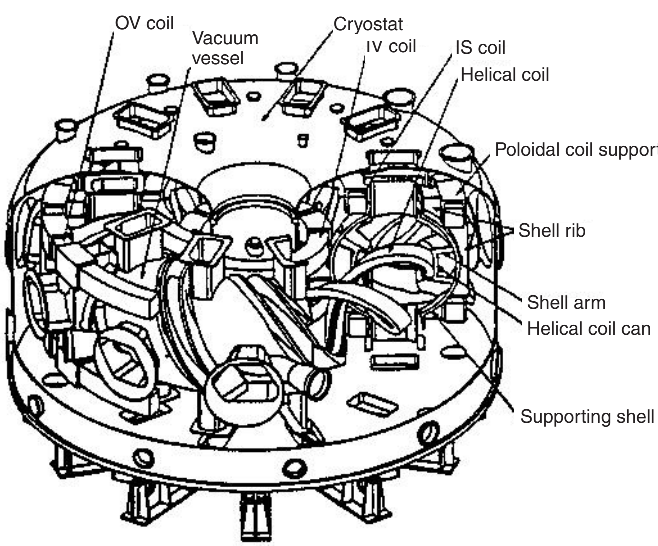


FIGURE 5-38 Schematic view of the large helical device (LHD).

TABLE 5-14 LHD Specifications

Parameter	Value
Major radius, m	3.9
Coil minor radius, m	0.975
Averaged plasma radius *l*, m	0.5 ~ 0.65
	2, 10
Magnetic field, T	3(4)
Helical coil current, MA	5.85(7.8)
LHe temperature, K	4.4(1.8)
Poloidal coil current	
Inner vertical coil, MA	5.0
Inner shaping coil, MA	−4.5
Outer vertical coil, MA	−4.5
LHe temperature, K	4.5
Plasma volume, m^3	20 ~ 30
Heating power, MW	40
Coil energy, GJ	0.9(1.6)
Refrigeration power, kW	9(~15)

The LHD began operation in fall 1998. The objectives of steady-state experiments are to (1) satisfy the necessary condition for an economical fusion reactor by ensuring self-ignition, possibly taking advantage of the currentless plasma; (2) increase the knowledge of plasma boundary phenomena, plasma-wall interactions, plasma heating (burning), fueling, particle and impurity control, and ash-exhausting and confinement improvement; and (3) advance fusion technology including high-heat-flux components, cooling techniques, and pumping systems.

LHD explores a new parameter regime involving long pulses >1000 s. The level of the steady (continuous-wavelength [CW]) power is 3 MW at least, with the temperature of more than a few thousand electronvolts in the medium-density region.

The helical divertor is one of the most important design features for control of severe plasma-wall interactions, and for heat and particle fluxes. It is also installed for improving the confinement performance of the plasma.

Elmo Bumpy Torus (EBT). The EBT concept is a toroidal array of simple magnetic mirrors. The promise of a steady-state, high-β reactor that operates at or near D-T ignition emerges from this combination of simple mirrors and toroidal geometry. The creation of an rf-generated, low-density, and energetic electron ring at each position between mirror coils (i.e., midplane location) is needed to stabilize the bulk toroidal plasma against well-known instabilities associated with simple mirror confinement.

Following early experiments on the concept at Oak Ridge National Laboratory, studies of this approach ceased in the early 1990s because of a combination of technical and funding difficulties.

Toroidal Bicusp (Tormac). Like the tokamak and the stellarator and Torsatron designs, the Tormac is a toroidal device that confines plasma on combined poloidal and toroidal magnetic fields. By opening the outer poloidal field regions, however, the Tormac creates an absolute minimum-B configuration that is MHD-stable for large aspect ratio and plasma β. The resulting toroidal line cusps support plasma on both closed field lines (i.e., high-β bulk plasma) and open field lines; confinement of the latter plasma is enhanced by mirroring effects in the sheath region that separates regions of open and closed field lines. Few experimental studies of this concept have been pursued since an initial series of work in the late 1980s.

Surface Magnetically Confined Systems (Surmac). The Surmac concept represents one example of a general class of multipole configurations in which electrical conductors are arrayed in either a linear or toroidal geometry to create a surface magnetic configuration with low magnetic field in the bulk plasma volume. Figure 5-39 illustrates a toroidal version of this high-β, steady-state system. The Surmac can operate with considerably reduced synchrotron radiation emanating from the bulk plasma, and, therefore, this concept appears to be particularly suitable for confining the high-temperature advanced-fuel plasmas.

Long-Pulsed Toroidal Systems. In terms of power density, relative simplicity, and symbiosis with the basic confinement scheme, ohmic dissipation of toroidal plasma currents represents a very efficient heating scheme. Two long-pulsed toroidal concepts are described that propose ohmic heating as the sole means to obtain an ignited thermonuclear plasma: the Riggatron and the reversed-field pinch (RFP). A variation of the RFP would use an external helical winding to achieve a more controllable rotational transform in a reversed-field state; this concept is called the *ohmically heated toroidal experiment* (OHTE).

The High-Field Ohmically Heated Tokamak (Riggatron Ignitor). Although in principle a tokamak, the Riggatron represents a sufficient change in engineering approach and "conventional" tokamak physics to warrant classification as an alternative concept (Coppi et al. 1992, Carpignano et al. 1995).

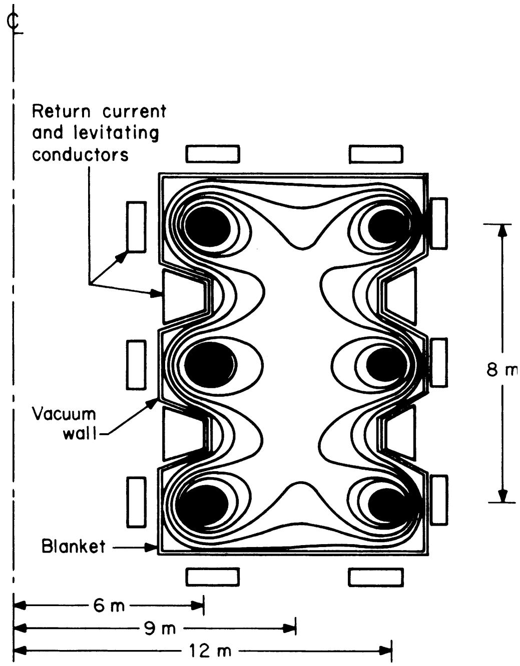


FIGURE 5-39 Schematic representation of the toroidal Surmac plasma and coil configuration.

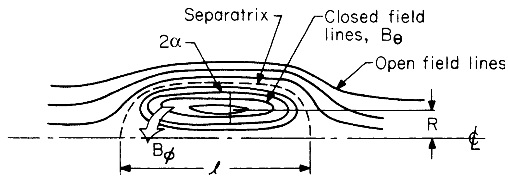
The combined use of high toroidal current density (8 MA/m^2) and high toroidal field (16 to 20 T) copper coils positioned near the first wall allows net energy production in a relatively short burn period from a high-β, ohmically heated system. The severe thermal-mechanical environment necessarily dictates an engineered short life. The plasma chamber and the D_2O-cooled copper magnets would be small because of the increased plasma density (2 to 3 $\times$ 10^{21} m^{-3}) and high β (0.15 to 0.25). The 6- to 10-MA Riggatron would generate 1 to 2 GW (thermal); the fusion neutron power is recovered in a fixed lithium blanket located outside the magnet system. Recovery of joule and neutron heating in the copper coils is also an essential element of the overall power balance.

The Ignitor experiment involves an advanced compact high-magnetic-field machine, somewhat similar to the Riggatron concept. It employs normal conductor cryogenic magnets with the goal of investigating D-T fusion burning plasmas. The design and construction of the experiment is being carried out by an industrial-university consortium in Italy in collaboration with MIT scientists.

Reversed-Field Pinch (RFP). The RFP is similar to a tokamak in that a toroidal axisymmetrical configuration is used to confine a plasma with toroidal current by a combination of poloidal and toroidal magnetic fields. Using a passive conduction shell, the RFP relaxes inherent constraints on the magnetic-field profiles for a tokamak such that the variation of the magnetic shear need not exhibit a minimum in a region enclosed by a first-wall conducting shell. By removing this constraint, the RFP can operate with a current density that is sufficient for ignition by ohmic heating, an unrestricted aspect ratio, a higher β value, and an appreciably lower magnetic field at the superconducting windings. Consideration has also been given to the possibility of developing a compact reactor based on the RFP concept but using copper coils. The compact RFP reactor would have many operational features similar to those described above for the Riggatron.

Pulsed Toroidal Systems. The early quest on the part of fusion reactor designers to attain the economic advantages of high-β operation simultaneously with the physics advantages of toroidal confinement led to concepts like the reference theta-pinch reactor (RTPR) and the high-β stellarator (HBS). It was generally found that the fast-pulsed nature of the RTPR (i.e., ~1- to 2-μs shock heating, 30-ms adiabatic compression, ~0.5-s burn time) resulted in technological problems that may outweigh the high-β ($>$0.8) advantages for that particular system. Additionally, the absence of MHD stability without fast feedback for the particular field configurations then under experimental investigation indicated other reactor-related problems for both the RTP and the HBS, although the latter was eventually proposed for steady-state operation. A more recent variation is the belt-shaped screw pinch reactor (BSPR). While it is still heated by a fast radial implosion, a toroidal bias magnetic field is applied to reduce the final values of β and thereby enhance stability.

Compact Toroids. The generic name *compact toroid* (CT) has recently been applied to the class of toroidal plasma configurations in which no magnetic coil or material walls extend through the torus.



PF (B_θ) versus TF (B_ϕ)	Larmor radius << a (MHD-like)	Larmor radius >> a (orbit-like)
$B_\theta \gg B_\phi \sim 0$	• FRC	• Astron-like ion compressors • e layers • p layers • REB-injected toroids
$B_\theta \cong B_\phi$	• Spheromak	

FIGURE 5-40 Diagram and classification of CT configuration where α is the minor radius, R is the major radius, B_θ is the poloidal field, B_θ is the toroidal field, and l is the separatrix length.

This closed-field plasmoid configuration is not new, as it was generated by a coaxial plasma gun a few decades ago. Interest in this configuration, as applied to a conceptual fusion reactor, however, began when the spheromak was proposed as a means to retain the developing physics base for tokamaks, while simultaneously shedding certain technological difficulties. Since the spheromak reactor was first proposed, the fusion community has identified the general area and potential of compact toroids; the spheromak is one element of the CT class of plasma configurations. Figure 5-40 summarizes the CT class of devices.

Although representing a great diversity of plasmoid formation, heating, and confinement schemes, the fundamental physics of particle and energy transport and stability and equilibrium are not well known for many subsets of the CT class. Two approaches that involve field reversal with smaller-orbit ions driving the plasma currents—the spheromak and the field-reversed theta pinch—are currently receiving the most study. The theta pinch-reversed field configuration, along with the field-reversed mirror, are normally classified as field-reversed configurations (FRCs).

The FRC and spheromak are closely related. A major difference is the lack of a torodial field component in the FRC, as its confinement properties derive solely from a polodial field.

FRC Reactor Studies. Various reactor studies (Momota et al. 1992, Choi et al. 1979, Miley 1988) have been performed to examine the potential advantages of FRCs. Both stationary designs and concepts where the plasma is passed through a linear burn chamber have been considered. The TRACT (triggered-reconnection adiabatically compressed toroid) reactor design for the stationary burn of a field-reversed theta pinch is a good example of these concepts. Table 5-15 also gives typical prototype reactor parameters for this ~1-Hz batch-burn system. Utilizing a longer burn period (~0.5 s) and modest magnetic fields (5.3 T), a hybrid superconducting (dc)/normal (ac) coil system would provide the required flux compression to achieve ignition in a plasmoid with an initial radius of 0.72 m. A first-wall copper coil cancels and subsequently reverses for a few milliseconds the field generated by an exoblanket superconducting coil, during which time a low-temperature

TABLE 5-15 Typical Parameters for TRACT, an FRC Reactor

Parameter	Value
Minor radius, m	0.14
Major radius, m	0.36
Length, m	1.88
Plasma density, $10^{20}/m^3$	28
Average β	0.77
Magnetic field, T	7
Pulsed energy, MJ	570
Burn time, s	0.5
Thermal power, MW	520
Net power, MWe	130

plasma is created. The superconducting flux is reestablished in two stages: a fast (shock) stage and a slower (adiabatic compression) stage. When the external field induced by the fast shock-heating power supply reaches a peak value, trigger coils are activated and field-line reconnection occurs. The resulting elongated FRC rapidly compresses axially to achieve an equilibrium; this compression provides the bulk of the heating. The resulting ~1.5-m-long plasmoid would attain ignition.

The TRACT approach leads to a relatively small pilot plant of moderate cost that may operate on the basis of near-term technology. A large commercial plant that distributes the power supply costs over several reactor modules benefits since the economy of scale for these systems predicts acceptable direct capital costs. An alternative pilot plant design, SAFFIRE, employed D-^{3}He fuel and neutral-beam injection for heating plus partial current drive. Pellet injection to maintain the desired radial density profile also contributed significant current drive in SAFFIRE. Reasonable costs for the commercial version would require cost savings by manufacturing of modular components in quantity.

The CTOR system designed at Los Alamos National Laboratory would use a field-reversed theta pinch to produce an FRC plasmoid external to the reactor that is subsequently translated through a linear burn chamber. Like SAFFIRE, in the CTOR design the plasmoid source and compressional heater are removed from the burn chamber to a less hostile environment. To minimize the technological requirements imposed by the plasmoid source and the associated pulsed power, a flared axial compressor would maintain the first-wall magnet coil close to the plasma for stability, while the translating plasmoid is adiabatically compressed to ignition prior to entering the linear burn chamber. Figure 5-41 shows a schematic view of this operation. Translation of the ignited plasmoid in the high-temperature burn chamber allows portions of the conducting shell that have not experienced flux diffusion to be continually "exposed." A nearly steady-state (thermal) operation of the first wall and blanket is possible by adjusting plasmoid speed and injection rate. Superconducting coils are located outside the blanket, conducting shell, and shield to provide a continuous bias field that is compressed between the conducting shell and the plasmoid; gross MHD stability would thereby be provided throughout the burn without requiring feedback stabilization. The plasmoid motion terminates in an end region where expansion directly converts internal plasma energy to electrical energy.

The most extensive FRC reactor design study is for a D-^{3}He system termed *Artemis.* This study, led by a design team at the National Institute for Fusion Science in Japan, was carried out for the purpose of examining its attractive characteristics and clarifying critical issues for commercial fusion reactors. The D-^{3}He-fueled FRC fusion reactor Artemis consists of a formation chamber, a burning chamber, and a pair of direct-energy converters, all of which are connected by magnetic lines of force (Fig. 5-42). An FRC plasma is produced at start-up by the conventional reverse-biased fast theta-pinch method in the formation chamber and then translated to the burning chamber.

A combination of deuterium NBI, fueling, and a slow magnetic compression in the burning chamber brings the volume, the temperature, and the number density of the plasma up to those for the D-^{3}He burning state. Most of the D-^{3}He fusion energy is carried by charged particles along the lines of force, which connect to a pair of direct-energy converters. A smaller fraction of the fusion energy is carried by neutrons and photons to the first wall of the burning chamber. This energy is converted to electricity by turbine generators.

The formation chamber is arranged symmetrically so as to reduce error fields. A fast-rising theta-pinch discharge with a one-turn voltage of 400 kV in the filling gas pressure of 0.05 Pa and bias field

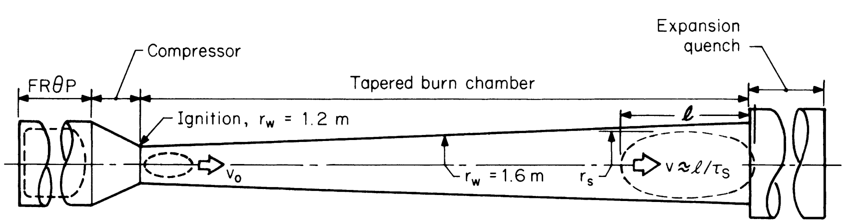


FIGURE 5-41 Schematic layout of a CTOR that would translate at an initial velocity ϑ_0 an FRC down a linear burn chamber of length *l*. The FRC would be formed externally to the reactor by an FRθP plasmoid source.

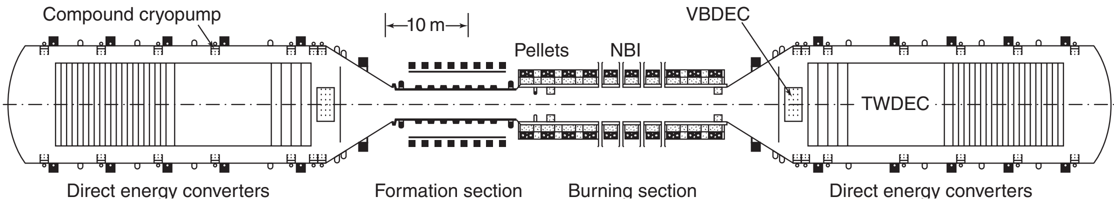


FIGURE 5-42 The D-^{3}He FRC reactor Artemis, composed of a formation chamber, a burning chamber, and direct-energy converters. The venetian blind convertor (VBDEC) handles the lower-energy background plasma and alpha particles, while the traveling-wave unit (TWDEC) handles the energetic protons.

of 0.035 T produces an initial FRC plasma (Table 5-16 column a). The plasma is then translated to the burning chamber because of an unbalanced cusp magnetic field. During the translation, the FRC conserves particle numbers of respective species, the total energy, the trapped flux, and its "intelligence." Plasma parameters obtained at the burning chamber are listed in Table 5-16, column b.

In the burning chamber, the FRC plasma is heated by means of energetic deuterium NBI with a maximum power of 100 MW. A slow magnetic compression is also applied. The injected particles form an ion beam current that serves as the seed current needed to sustain or increase the trapped magnetic flux of an FRC. The plasma evolves in its volume, the density, and the temperature, forming a burning plasma. During the evolution, the ratio of the current carried by the energetic beam particles to the total plasma current is high enough to stabilize an FRC against macroscopic modes. The set of plasma parameters obtained in this way is tabulated in Table 5-16, column c.

In a D-^{3}He burning FRC plasma, the current drive due to the preferential trapping of fusion protons, assisted by a small amount of external beam injection, is sufficient to sustain an equilibrium in a steady burning state.

The concept of a traveling-wave direct-energy converter (TWDEC) was developed to extract the fusion energy carried by 14.7-MeV fusion protons. The technologies needed for this development are conventional. The lifetimes of existing structural ferritic steels are as long as the total reactor life of the Artemis, and no development of new materials is necessary. Fuel injection is accomplished by oscillating the FRC axially to "engulf" fuel pellets dropped into its path. Operation of large high-field magnetic coils is not required. Thus, the bases of technologies employed for Artemis design are conventional and have considerable technological flexibility that enables the design to accommodate a range of unknown parameters.

The estimated cost of electricity (COE) from Artemis is ~30 mill/kWh, which is low compared with a tokamak reactor or even an LWR. The COE depends only weakly on the fuel cost; thus, uncertainties in the cost of ^{3}He are not crucial. The high energy-conversion efficiency (~62% overall) obtained with the TWDEC and the modest cost of that unit are key factors in the competitive COE

TABLE 5-16 Plasma Parameters at Various "Stations"

	Phase		
	a	b	c
Plasma radius, m	0.7	1.0	1.12
Plasma length, m	4.8	14.8	17.0
Plasma temperature, keV	1.0	1.0	87.5
Electron density ($\times 10^{20}/m^3$)	4.1	0.62	6.6
Trapped flux, Wb	0.086	0.086	3.66
External magnetic field, T	0.56	0.22	6.7
s value	5.9	2.7	9.2
Plasma beta value	0.92	0.83	0.90

Key: a—after formation of the FRC; b—after translation of the FRC; c—steady burning state.

obtained for Artemis. Low radioactivity, low tritium yields, and the inherent safety of the power plant should be readily acceptable socially and ecologically, representing an imporant advantage of this type of power plant.

Studies on issues such as stabilization of macroscopic modes by means of energetic particles must still be carried out, but the general characteristics of a D-^{3}He-fueled FRC fusion power plant offer a very attractive prospect for energy development for the twenty-first century.

A key issue in the use of a D-^{3}He reactor such as Artemis is the source of ^{3}He. In the design study, lunar mining was assumed and the corresponding cost factored into the COE cited. An alternative approach is to breed ^{3}He, either using a semicatalyzed D-D reactor or from production of excess tritium and subsequent decay to ^{3}He.

FRC Experiments: LSX. Reviews (Haruhiko et al. 1995, Hoffman 1996, Hoffman et al. 1993, Momota et al. 1992) of the research status of the FRC discuss small-scale experiments under way at the following laboratories worldwide: BN (TRINITI, Moscow), TL (TRINITI), TOR (TRINITI), NUCTE-3 (Nihon University), FIX (Osaka University), MRX (PPPL), STX (University of Washington), FIREX (Cornell University), and rotomak (Flinders University). Notable achievements include formation by a theta pinch and by counterhelicity injection, stabilization of the rotational mode, detection of global internal modes, tilting-mode theory, translation and acceleration, identification of transport anomalies, and reduction of convective thermal losses. Immediate issues include the need to demonstrate a long-lived, high-quality FRC, improved theories for global stability land kinetic physics, the development of a reactor-relevant start-up method and current drive, and a better understanding of the plasma energy-transport mechanism. Additional long-term issues include studies of flow stabilization and kinetic stabilization effects, rotating-magnetic-field current drive, and the concept of a traveling-wave direct-energy converter (TWDEC). FRC experiments have undergone a 5-year hiatus in the United States because of the 1990 DOE decision to halt alternative confinement fusion research in favor of concentrated tokamak development. The largest FRC facility, the Large *s* Experiment (LSX), was built during 1986 to 1990, but operated for only 1 year, after which it was converted to studies of plasmoid fueling tokamaks. Its parameters are shown in Table 5-17. A modified version of that facility was reinaugurated in 1998.

The LSX was constructed to explore the stability of FRCs as *s* (device radius/ion gyroradius) was increased beyond the value of 2 (produced in previous small experiments), where stabilizing kinetic effects are theoretically predicted to diminish. Stable, well-confined FRCs were produced with *s* values up to 4, with indications that higher-*s* FRCs would be stable if they could be successfully formed using the theta-pinch techniques. However, before the LSX program was halted, it became clear that theta-pinch formation was limited to flux values in the tens of milliwebers range, while webers of flux are required for a reactor. Thus, flux buildup represents a key area for future research.

The particle-confinement times measured on LSX and other FRC experiments indicate a strong correlation with the parameter $rs/\sqrt{\rho_i}$, where rs is the separatrix radius (see Fig. 5-40) and ρ_i is the gyroradius. The general form implies a confinement time scaling ($\tau_N = x_s\, a^2/D_\perp$) with a basic particle diffusivity [$D_\perp = 2/n^{1/2}$ (10^{21} m^2) m^2/s], where a is the minor radius of the toroid (Fig. 5-40) and n is the averaged plasma density. The factor x_s occurs here as a result of the increase in flux (lower beta) for a given radius FRC, and the dependence on density just reflects the $\rho_i \propto 1/n^{1/2}$ relationship for a high β plasma.

This empirical scaling is unfavorable for low-density operation, and thus low-power, steady-state operation. However, there is evidence from recent Japanese experiments (Haruhiko et al. 1995) that an improved low-density confinement mode may exist. Confinement may also improve when current sustainment techniques are applied. Then, the toroidal plasma current is carried either by large-orbit ions or gyrocenter drifts, rather than by simple diamagnetism.

TABLE 5-17 LSX Parameters

Property	LSX experiment
Coil diameter/length, m	0.9/5
Separatrix diameter/length, m	0.45/4
$x_s/\langle\beta\rangle$	0.5/0.87
External magnetic field, T	0.4
Poloidal flux, Wb	0.010
Plasma density, m^{-3}	1.2×10^{21}
Plasma temperature, keV	0.15
Ion gyroradius, cm	0.44
Kinetic parameter *s*	4

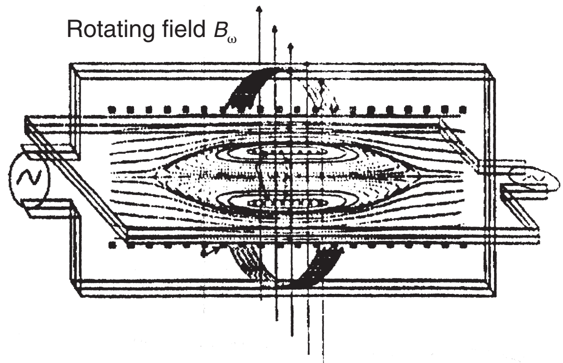


FIGURE 5-43 RMF current-drive schematic.

In order to build a steady-state FRC reactor it is essential to develop some form of flux buildup and current drive. The Artemis reactor design study described earlier assumed neutral beams resulting in large ion orbits (Momota et al. 1992). Large-orbit ions are theoretically effective in stabilizing compact toroids, but the technology is expensive and the large azimuthal ion-ring momentum may introduce ring-plasma instabilities. A simpler, but unproven current drive method for hot plasmas would be to directly drive the electrons using rotating magnetic fields (RMFs). Such a technique, sketched on Fig. 5-43, has been demonstrated on small-scale rotomak experiments.

The two antenna coils in Figure 5-43 are driven 90° out of phase to produce a rotating field $B_r = B_\omega \cos(\omega t - \theta), B_\theta = B_\omega \sin(\omega t - \theta) E_z = r\omega B_r$. If $\omega_{ci} < \omega << \omega_{ce}$, then the electrons, but not the ions, will be dragged along with the rotating field. Here ω_{ci} and ω_{ce} are the ion- and electron-cyclotron frequencies, respectively. The perpendicular B_ω field will prevent an axial current from flowing and shielding the plasma interior from the rotating field.

Spheromak Reactor Concepts. The spheromak lends itself to a reactor embodiment (Fowler et al. 1994, Hagenson and Krakowski 1985, Hooper and Fowler 1996, and Moir 1996) that shares many features with the FRC. However, fewer specific studies have been reported for the spheromak. Still it has been shown to potentially offer a comparatively simple fusion reactor with relatively low cost of electricity. The magnetic fields are generated primarily by plasma currents except for a vertical field, required to support the hoop stress, which is generated by a set of purely solenoidal coils. Current is driven by an external coaxial gun that generates linked toroidal and poloidal magnetic fluxes (helicity) that drive current in the core of the spheromak through a magnetic dynamo. The plasma-core current may return to the gun or be collected at the other end of the magnetic separatrix. It is anticipated that the ohmic heating from the plasma current will be sufficiently strong such that no auxiliary heating source will be required to reach reactor conditions.

Figure 5-44 shows the proposed magnetic-field configuration for a spheromak reactor. The fusion plasma is confined within a magnetic separatrix with toroidal, magnetic flux surfaces; thus, the geometry shown has two *X* points. Helicity injection is from a coaxial, electrostatic gun shown located above the upper *X* point; power flowing from the confinement region is diverted into the upper and lower coaxial regions, which act as divertors. It may be possible to unbalance the magnetic configuration so that most of the energy and particles flow to the lower region, thus separating the divertor from the gun, and potentially improving the power and particle handling characteristics of the reactor.

The blanket, located in the annular region surrounding the plasma, is envisioned to consist of liquid lithium or lithium salts, perhaps flowing vertically between the wall and the plasma; solid or other configurations are also possible. If ohmic ignition occurs, the spheromak will allow a blanket without radial port penetrations. Thus, it has been suggested that the simplicity of the configuration may make a liquid/NaK "potboiler" type of reactor possible.

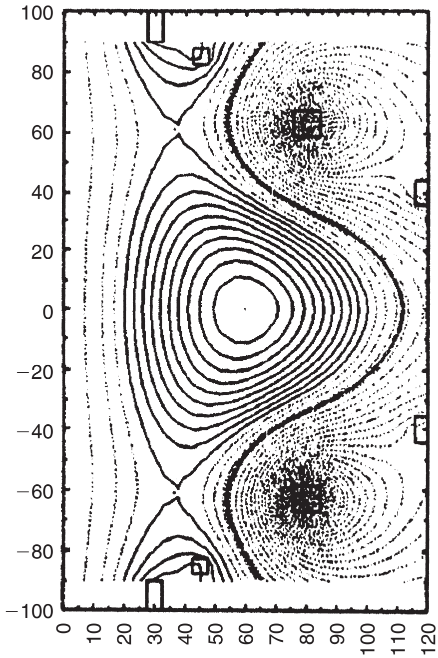


FIGURE 5-44 Flux surface geometry for the spheromak reactor concept. Helicity is injected by flowing current along the open flux surfaces from the upper to the lower divertor regions.

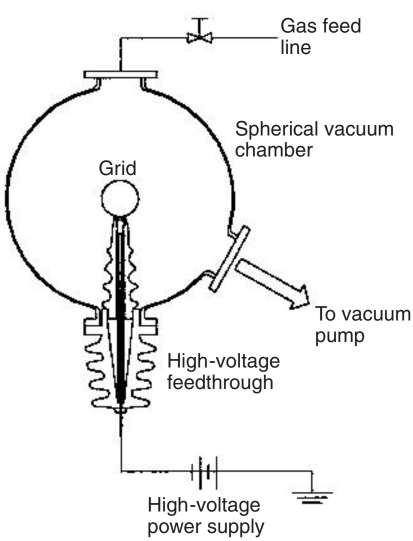


FIGURE 5-45 Cross section showing construction of the single-grid spherical IEC.

5.4.9 Inertial Electrostatic Confinement

Instead of using magnetic fields for confinement, electrostatic fields can be employed (Miley 1997, Kulcinski et al. 1997, Miley et al. 1997, Sved 1997, Gu and Miley 1995, Miley et al. 1998). This requires a dynamic plasma, for instance, one where the inertia of high-energy ions becomes significant since, as proved by Earnshaw, pure electrostatic fields cannot be used alone to confine a plasma. Thus, electrostatic fields have generally been viewed as a way to "plug" plasma leakage routes in magnetic confinement systems, which has been widely considered (e.g., see the preceding discussion of the tandem mirror), and studies of electrostatic confinement have been limited. The best known concept, termed *inertial electrostatic confinement* (IEC), uses the inertia associated with recirculating ions to overcome Earnshaw's theorem. This concept traces back to Philo Farnsworth, the U.S. inventor of electronic television, who proposed and patented aspects of the concept in the 1960s. While some continuing research has been done ever since then, the first aggressive small-scale experimental projects were set up at the University of Illinois, Los Alamos Scientific Laboratory, University of Wisconsin, MC^2 Corporation, and the University of Kyoto.

A significant result of these efforts is that Daimler-Benz Aerospace Corporation initiated commercial production of a version of the University of Illinois unit. The commercial IEC unit is for use as a portable 2.5-MeV neutron source (D-D fusion) for industrial neutron activation analysis (NAA) and neutron tomography. While operating with a net power input well below energy breakeven ($Q \sim 10^{-6}$), this development represents the first commercial use of a confined fusing plasma.

The University of Illinois experiment is representative of present IEC research and will be briefly described here. For simplicity, it replaces the ion guns used in early Farnsworth-type experiments with a grid-produced plasma discharge. Operation employs the "star" mode, where ion beams formed in the discharge pass through grid openings, minimizing grid sputtering and erosion.

Both spherical and cylindrical versions have been developed. In the spherical design, illustrated in Fig. 5-45, the transparent grid biased at −60 kV acts as a cathode relative to the grounded vacuum vessel wall.

When a deuterium gas (or deuterium plus tritium or ^3He mixtures) discharge is used, D^+ ions are extracted from the plasma by the cathode grid, accelerated, and focused in the center where D-D fusion reactions occur. The grid provides recirculation of the D^+ ions, increasing the power efficiency. At high currents, a potential structure develops in the nonneutral plasma, creating virtual electrodes that further enhance ion containment and recirculation (Fig. 5-46).

Steady-state spherical IEC units typically produce 10^7 D-D n/s, while advanced pulsed versions extend to 10^{11} neutrons per second (n/s) (10^{13} n/s for DT).

The cylindrical version of the IEC (Fig. 5-47) uses a hollow electrode design where ions recirculate between two outer anodes producing a line-like neutron source within the hollow cathode region. This unit is most useful for applications requiring broad-area extended sources, such as luggage inspection, and NAA of materials, such as ores, on a conveyer belt.

Neutron yields from the cylindrical design are comparable to those from the spherical IEC, but the lack of three-dimensional virtual electrode formation limits its ultimate fusion rate.

The unique feature of the IEC is that it develops a fusion-grade plasma in a small volume (a few cubic centimeters) while avoiding bulky magnets, injectors, and the like. In addition, the star mode creates ion beams that reduce grid bombardment, which, combined with use of a "damage free" plasma target (the dense core of Fig. 5-46) provides long-lifetime units. Thus, the commercial IEC NAA units provide a first step in what could be a progression of uses for fusion-grade plasmas prior to actual fusion power plants.

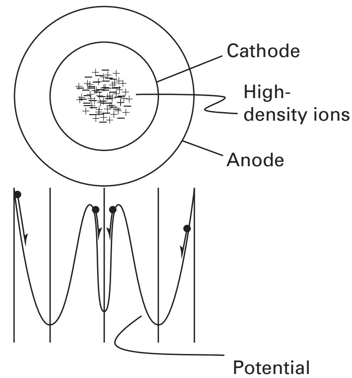


FIGURE 5-46 The potential structure in the dense plasma core of an IEC. The central "well" traps ions, providing efficient local recirculation.

In addition to NAA, the IEC has been proposed as a tunable x-ray source and as a plasma thruster for space applications. These applications generally require higher reaction rates, and present research is focused on that objective, largely via use of higher ion currents. If successful, the scale-up to a near-breakeven unit could be considered for neutron-irradiation applications such as neutron-damage studies. Development issues include the stability of the potential structure illustrated in Fig. 5-46 at high currents and technology such as grid survivability. Grid scale-up requires development of very efficient active cooling methods, or use of an alternative approach to eliminate grids. For example, a recent concept proposed by R. W. Bussard replaces the grid with a unique high-order cusp magnetic field.

5.4.9 Inertial Fusion Energy and Concepts

In the inertial confinement scheme, the surface of a small spherical target of a solid material containing deuterium-tritium fuel is symmetrically illuminated by very high energy laser or charged particle

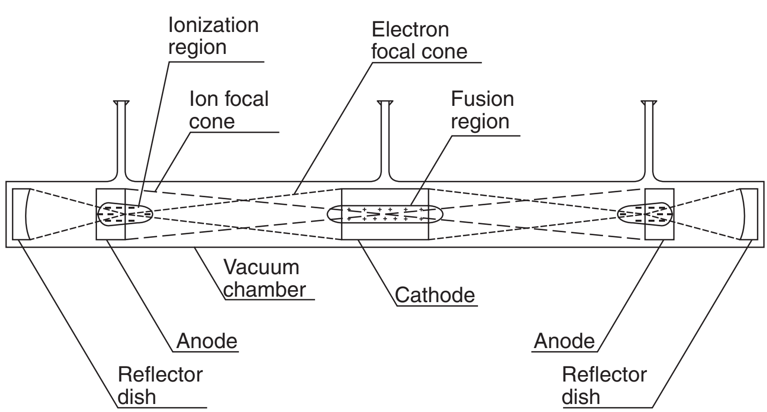


FIGURE 5-47 Cross section of the cylindrical version of the IEC. The ion-electron optics are designed to provide a central beam through the hollow cathode.

beams. The density of the nuclei in the target prior to illumination is typically ~10^{22} ions/cm^3. A thermal compression, caused by ablation as the beam energy impinges on the target, causes a further increase in density of up to 10^4 times, producing a density at the center of the pellet approaching 10^{26} ions/cm^3. At this density, a Lawson-type breakeven density-confinement time product of 10^{16} s/cm^3 can be achieved in a confinement time of only 1 ns (10^{-9} s) or less, which is achieved by the inertia of the fuel itself. If a central pocket of DT gas is left in the center of the target, as shown in Fig. 5-44, it will be adiabatically heated during compression to a temperature of 10^8°C, initiating a fusion burn. These density-confinement, time-temperature conditions are adequate for energy breakeven. Thus, the process is essentially one of a thermally induced implosion. The total laser energy which must be applied to the target to achieve useful fusion gains is estimated to be of the order of 5 MJ. With a target gain of 100, this would give a yield of 500 MJ. A pulse rate of 2 Hz would produce 1000 MW_t.

In terms of power plant design, the most important feature of inertial fusion is that the confinement of the plasma is decoupled from the functions of the reaction chamber, wall, and blanket. This is possible because the physics of energy absorption, implosion, ignition, and burn of this plasma are physically separable from the chamber conditions and structural requirements.

In inertial fusion power plant chamber design, the degree of freedom allowed by this separability is exploited by placing fluids (gases or liquids) or fields (magnetic deflection) inside the chamber to protect the wall from ablative material loss due to the energy deposition of the short-range fusion radiation (25% to 35% in soft x-rays and debris) and, in the case of thick liquid-metal walls, to provide protection from radiation damage due to deeply penetrating x-rays and neutrons. Similarly, the final optics of the driver can be protected by placing the final mirrors tens of meters away and protecting them by a gas, or in the case of an ion beam driver, by closing a rotating shutter in front of the ion beam ports to block the plasma debris.

The use of protective fluids, combined with the pulsed nature of the energy release, leads to new engineering constraints and techniques that are radically different from those of magnetic fusion. Understanding and manipulating the dynamic response of fluids and structures become of primary importance, and such factors as new materials development to resist neutron damage take less emphasis.

The HYLIFE System. The most detailed calculations and conceptual design for an inertial fusion power plant were performed at LLNL, in conjunction with a team of university and industrial contractors. The high-yield lithium injection fusion energy (HYLIFE) chamber shown in Fig. 5-48 is designed to operate with yields of a few thousands megajoules at rates of a few hertz. A lithium energy-conversion blanket, consisting of a dense array of 20-cm-diameter jets, is continuously injected into the chamber. This provides an effective blanket thickness of 1 m between the fusion pellet and the inner steel wall. The 14-MeV neutron flux is reduced by a factor of 200 by this lithium blanket, allowing a chamber with a 5-m-radius to operate for 21 full-power years (30 years at 70% availability), at an integrated neutron energy flux of 0.3 MW/m^2. Since the flux in the wall without lithium protection would be 5.7 MW/m^2 (including pellet effects), the power density within the reactor vessel is very high, approaching that of a fission reactor. A common low-alloy ferritic steel (2.25 Cr-1 Mo) is used throughout. The tritium-breeding ratio is controllable between 1.0 and 1.7 by adjusting one *G*/*i* enrichment in the lithium.

The chamber contains a hexagonal array (50% packing fraction) of 20-cm-diameter lithium jets, which gives an effective shock isolation from the effects of the fusion pulse. Since the hot gas merely blows through the array of jets, this configuration minimizes the wall stress due to the impact of lithium accelerated by high-pressure blowoff gas, caused by x-ray and debris energy deposition. A substantial fraction of the kinetic energy of expansion of fluid, resulting from the neutron absorption, is dissipated in the liquid-liquid interactions among colliding jets. Finally, the enormous surface area of the fluid acts as an effective condensation pump for the lithium vapor.

The HYLIFE power plant requires 16 pumps to inject the liquid lithium into the 5-m-radius chamber. The power necessary for the lithium recirculation pumping is 20 MW (electric), only 1.6% of the gross electric power, if mechanical pumps are used. The laser mirrors are located at the far ends of the containment building, 60 m from the microexplosion. They are protected from the soft x-rays and debris by 0.25 torr of xenon in a 30-m section of the passageway. The neutron flux on the

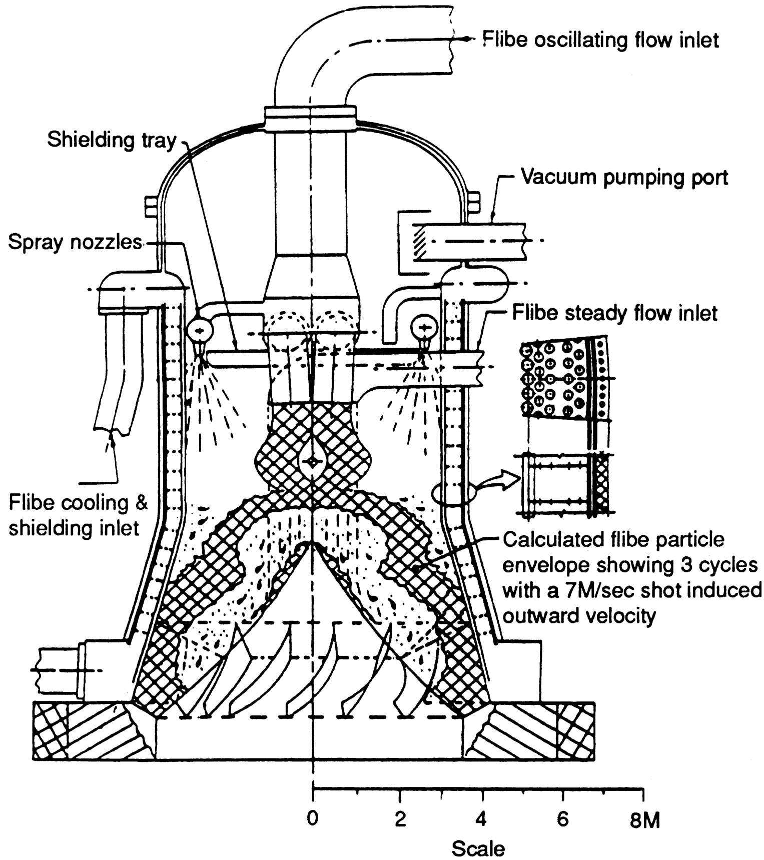


FIGURE 5-48 HYLIFE-II, oscillating flow combines wall protection and driver-beam access.

mirrors is 0.042 MW/m^2, low enough to assure substrate lifetimes of over a year. The power plant characteristics are summarized in Table 5-18.

HYLIFE-II. The HYLIFE-II reaction chamber, shown in Fig. 5-48, employed an oscillating flow scheme to permit beam injection while still protecting the chamber wall. HYLIFE-II avoids the fire hazard of lithium by using a molten salt composed of fluorine, lithium, and beryllium (LI_2BeF_4) called Flibe. Access for heavy-ion beams is provided.

A high repetition rate is employed in HYLIFE-II, requiring faster reestablishment of the jets after a shot. This is accomplished in part by decreasing the jet fall height and increasing the jet flow velocity. In addition, there is liquid splash that must be forcibly cleared because gravity is too slow at repetition rates higher than 1 Hz. Splash removal is accomplished in the central region by oscillating jet flows.

Flibe is compatible with 316-stainless steel or even better with Hastelloy N at a much higher temperature than lithium (923 vs. 770 K). Its heat-transfer properties, while different from lithium, are still suitable to remove heat and serve to protect the permanent structure from neutron damage and blast. Flibe, unlike lithium, dissociates. However, recombination is rapid enough to avoid affecting condensation; hence this does not limit the repetition rate. If the salt is maintained in a reduced state

TABLE 5-18 Summary of HYLIFE Power Plant Characteristics

System parameters	
Fusion power, MW	2700
Net electric power, MW (electric)	1004
Net system efficiency, %	32
Tritium breeding ratio	1.0 to 1.7
Laser and pellet parameters	
Beam energy, MJ	4.5
Pellet gain (energy multiplication)	400
Yield, MJ	1800
Repetition rate, Hz	1.5
Laser efficiency, %	5
Laser power consumption, MW (electric)	135
Fusion energy gain	20
Fusion chamber	
Radius, m	5
Height, m	8
Material	$2^1/_4$ Cr-1 Mo
Midplane neutron flux, MW/m^2 (with lithium)	0.68
Lithium array geometry	
Number of jets	175
Midplane jet diameter, m	0.2
Midplane packing fraction	0.57
Effective array thickness, m	0.47
Flow parameters	
Midplane jet velocity, m/s	13.3
Array flow rate, m^3/s	72.2
Total lithium pumping power, MW (electric)	26.0
Total lithium inventory, m^3	1850
Lithium outlet temperature, °C	500
Lithium temperature rise per pulse, °C	18

Source: Krakowski, R. A., Glancy, G. E., and Dabiri, Ali E., The technology of compact fusion reactor concepts. *Nuclear Technology/Fusion.* 1983, vol. 4, p. 342.

by keeping the mole fraction of tritium fluoride (TF) low, corrosion from fluorine compounds formed during the evaporation process should be small. The salt is reduced by contacting it with beryllium or by use of a CeF_3/CeF_4 buffering agent.

To achieve a high repetition rate and short fall distance with splash clearing, the jet nozzles oscillate horizontally, as shown in Fig. 5-48.

"Cool" Flibe at 873 K is injected in a spray of droplets in the vicinity of the beam paths. This injected spray can provide a large condensation area. Thus, it is not necessary for the explosion itself to form a large number of small droplets for seeding the condensation process as is sometimes assumed.

Practically, all of the tritium gas emitted by exploding targets will be removed by the vacuum pumping system, but almost none of the tritium bred in the Flibe will diffuse out of the Flibe droplets while in the chamber. The reason is that tritium has a low solubility in Flibe so it tries to diffuse out of and not into the droplets. A vacuum disengager has been designed to remove the tritium from the molten salt coolant (see Fig. 5-49).

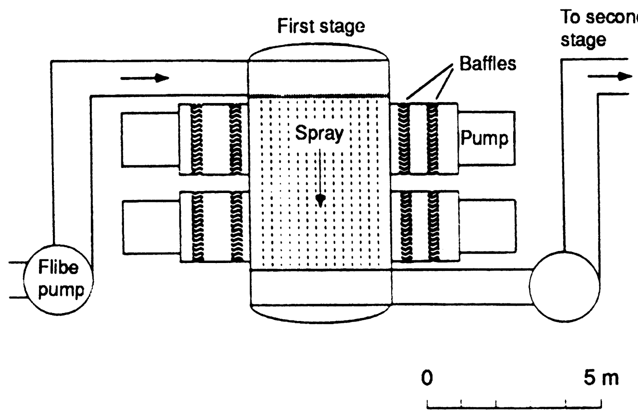


FIGURE 5-49 The vacuum disengager removes tritium from small Flibe droplets falling 5 m.

An outstanding feature of the HYLIFE-II reactor is its favorable safety characteristics. These include

Off-site dose from severe accident less than 2 Sv (200 rem) for passive safety

No N-stamp requirement for most components, thus requiring less than 0.25 Sv (25 rem) off-site dose

Working area dose rate less than 50 mSv/h (5 mrem/h) for a low occupational risk

Dose from routine atmospheric effluents less than 50 mSv/year (5 mrem/year)

Some of the key issues include verifying splash removal techniques, tritium removal effectiveness, condensation phenomena, heavy-ion accelerator technology, cost reduction, and beam propagation. Most importantly, to be competitive with future coal and light-water-reactor (LWR) nuclear power, the cost of electricity needs to be reduced by a factor of somewhat less than 2.

OFE Inertial Fusion Reactor Studies. There have been a number of ICF reactor studies such as HYLIFE with varying degrees of completeness (Miley and Campbell 1997). Recently, the Office of Fusion Energy (OFE) commissioned two studies to take advantage of the database through 1990, and we will briefly consider them. These studies take 1 GWe as the appropriate output electrical energy (a large power plant).

The OFE reactor studies have considered both heavy-ion and KrF laser drivers. Both approaches to ICF, namely, direct and indirect drive (qualitatively illustrated in Figs. 5-50 and 5-51), are considered. (Other studies are advancing the diode-pumped solid-state laser discussed earlier. A light-ion-beam driver study is also being advanced.) However, with at least thousands of active components, it is difficult to accurately assess the reliability and true costs of these conceived drivers. Figures 5-52 and 5-53 illustrate delivery of the driver energy in the Prometheus-H (heavy ion) and SOMBRERO (KrF laser) concepts that are taken as typical. Prometheus-H has a single accelerator, but 18 beamlets are generated, stored, delivered, and focused from two sides onto the indirectly driven target.

The SOMBRERO layout illustrates some of the issues for laser direct drive. Note that the building has a 55-m radius. The size of the building is dictated by the need to place laser optics far from the neutron source.

The ability to line up and consistently hit a target is a key issue. Figure 5-54 illustrates a delivery concept for the SOMBRERO case. The targets must be delivered accurately and, particularly in the case of indirect drive, they must also be precisely aligned. The OSIRIS reactor chamber concept (Fig. 5-55) has highly attractive features. The walls are wetted and continuously renewing. Liquid Flibe (500 to 600°C) is the wetting agent, the heat-removal agent, and the tritium-breeding agent.

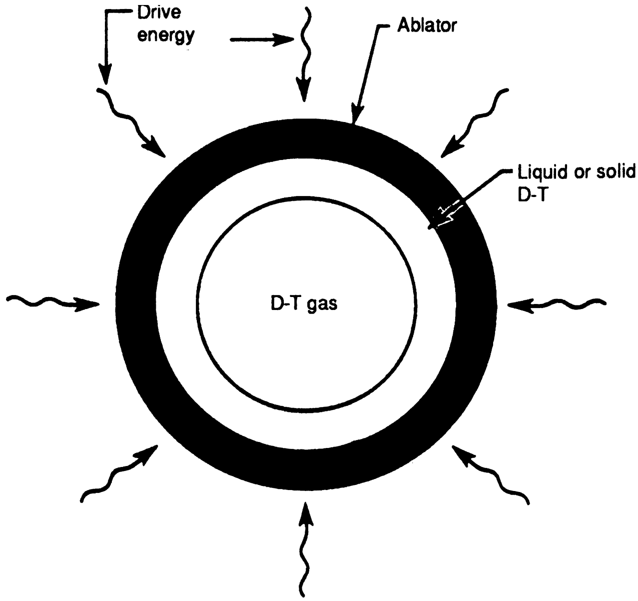


FIGURE 5-50 The basic concept of inertial fusion uses uniformly distributed radiation to drive the compression of a capsule containing deuterium and tritium. Direct drive applies the incident energy directly from a laser or ion beam.

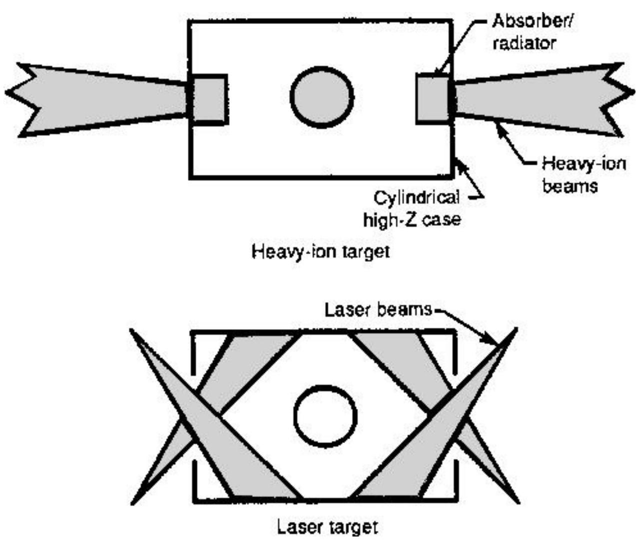


FIGURE 5-51 For indirect-drive targets, the incident ion beam or laser energy is focused onto a material surface where secondary radiation is generated, resulting in uniform radiation incident on the central, spherical capsule.

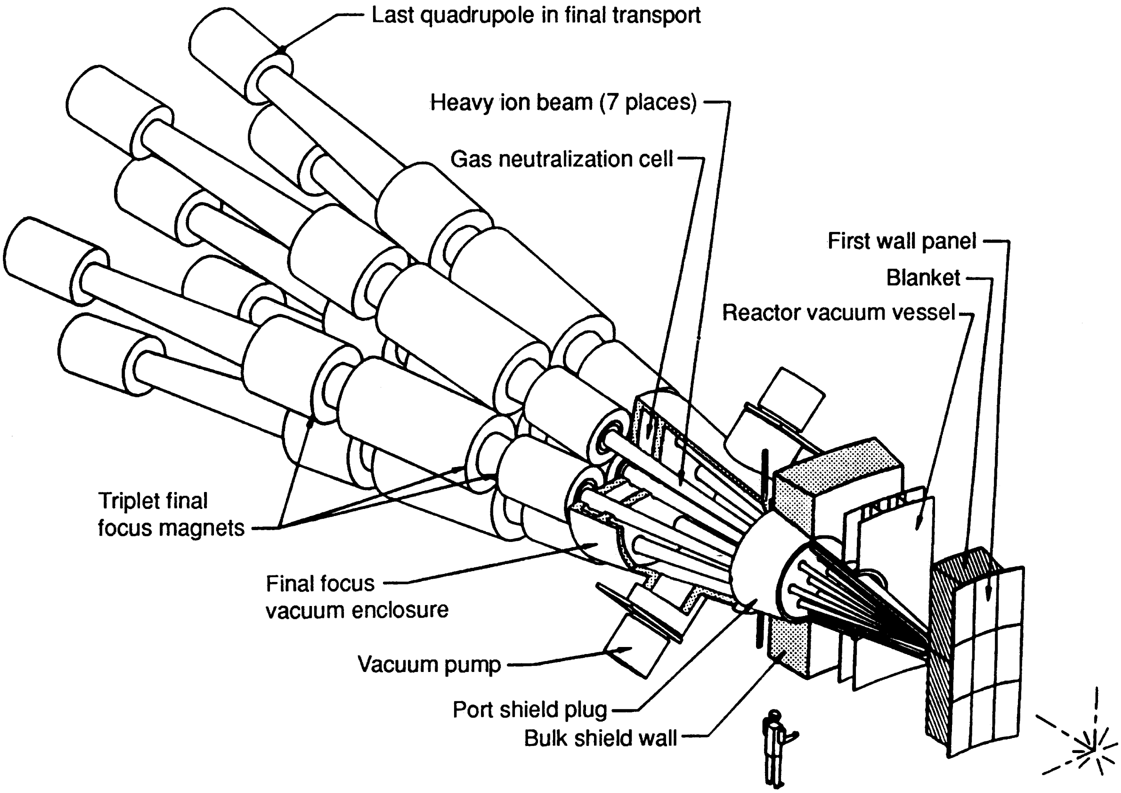


FIGURE 5-52 Side view of final focusing arrangement for one side of heavy-ion-driven Prometheus-H. (*McDonnell-Douglas.*)

Fast Ignitor. In addition to research on direct- and indirect-drive targets for inertial confinement fusion (ICF), worldwide interest has developed the concept of "fast ignition" proposed in 1994 by researchers at LLNL (Tabak 1994). In this concept, the target is first "hit" by the main driver beam which compresses it to a high density. Almost simultaneously, it would be "hit" by a second beam from a Petawatt laser, which uses chirped-pulse amplification to achieve very high power, with short time pulses through a thousand-fold pulse compression in time (Perry et al. 1996).

When focused onto a solid target, the Petawatt laser pulse can generate a forward-traveling beam of high-energy electrons. The fast-ignitor concept is to use this electron beam to "bore a hole" through the outer region of a precompressed ICF target, igniting a small volume of fusion fuel near the center (equivalent to "spark ignition" by converging shock waves created by a time-shaped laser pulse in normal direct-drive ICF).

Fast ignition can reduce the driver-beam energy by 10 times or more for inertial fusion energy (IFE) by reducing the implosion velocity and energy for a given fuel compression ratio (Logan and Meier 1998). For any type of driver that can deliver the ignition energy fast enough, fast ignition increases the target gain compared to conventional targets. By reducing the fuel compression energy, velocity, and implosion convergence ratio, and by tolerating mixing in the DT core, fast ignition also reduces the degree of symmetry required both on the driver beam illumination as well as on the smoothness of capsule ablators and cryogenic DT fuel layers. In principle, these advantages of fast ignition can accrue to any type of driver used to compress the fuel, and to any driver type used to ignite the fuel after compression; that is, ion accelerators or lasers might be used for either or both

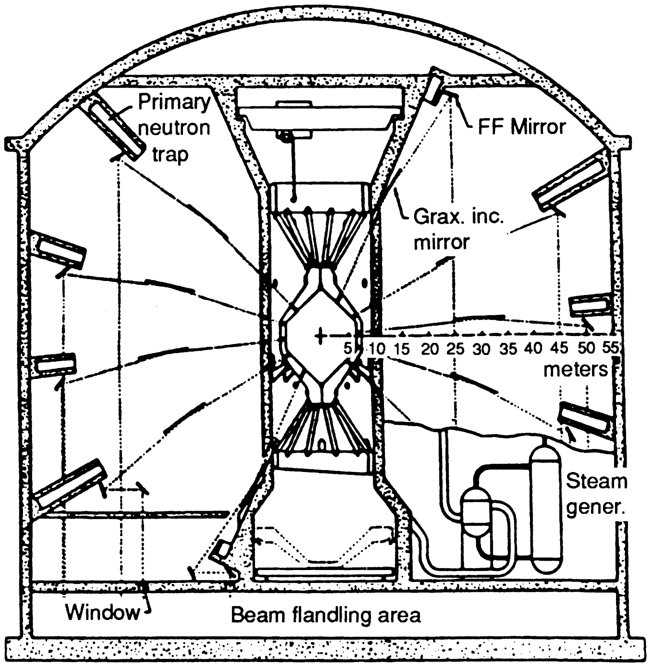


FIGURE 5-53 Side view of full reactor hall for the SOMBRERO KrF laser-driven concept. (*W. J. Schafer.*)

functions. With high peak beam power and small focal spots, fast ignition is more demanding of the driver than is fuel compression, despite lower-energy input required for ignition. Experiments with a Petawatt laser have demonstrated adequate conversion efficiency of light into hot electrons with the range needed to ignite DT cores in future experiments (Perry et al. 1996). Further experiments are required to determine if the laser light and ht electrons can channel deeply enough into the corona plasma of a preimploded capsule to heat a small portion of a >1000-fold liquid density compressed DT core to ignition. Ion-driven fast ignition, compared to laser-driven fast ignition, has the advantage of direct (*dE*/*dx*) deposition of beam energy to the DT, eliminating inefficiencies for conversion

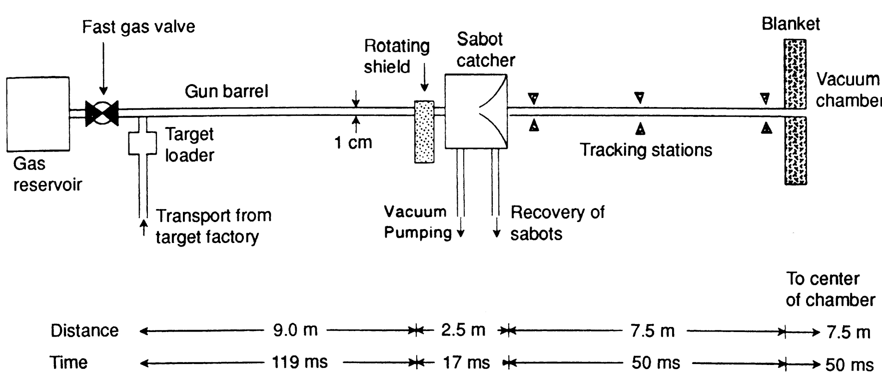


FIGURE 5-54 Target injection system for the SOMBRERO ICF reactor design. (*W. J. Schafer.*)

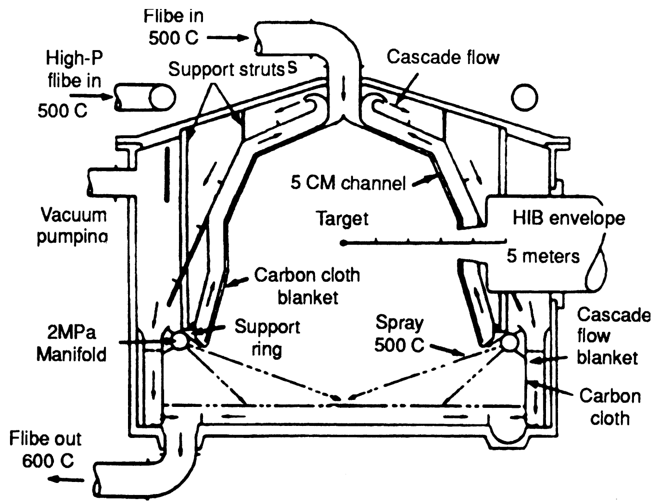


FIGURE 5-55 Side view of OSIRIS, heavy-ion-driven concept. The Flibe flows in through the narrow, 5-cm channel and soaks through the carbon-cloth blankets, providing a self-renewing wetted wall. The Flibe then circulates through the blanket, absorbing the neutrons for thermal energy and tritium breeding. (*W. J. Schafer.*)

into hot electrons. On the other hand, ion accelerators have not yet been demonstrated that can produce sufficient beam quality for high pulse compression ratios and small focal spot sizes for fast ignition. Accordingly, it is the accelerator and beam focusing, not the target coupling physics, that poses the main challenge to ion-driven fast ignition.

ICF Laser Facilities NIF. The NOVA laser at the Lawrence Livermore National Laboratory (LLNL) in California has served for years as the world's most powerful laser for inertial confinement fusion (ICF) research. It was shut down to make way for the National Ignition Facility (NIF). This facility has as its ultimate goal the demonstration of ignition and high-gain operation of ICF targets. To accomplish this, the NIF is designed to deliver over 2 MJ of 0.35-μm light to indirect-drive targets with a peak power of 500 to 1000 TW in a variety of pulse shapes.

For ICF, ignition is defined where the fusion alphas deposit sufficient energy locally to double the burn temperature in the hot spot of the fuel. This will occur in a DT target if the core is compressed to $\rho r > 0.3$ g/cm^2 and a preignition temperature of 5 keV. Then a thermonuclear burn will propagate into the remaining cold fuel, which is nearly isentropically compressed to high density. For a reactor, an energy gain of 100 or more would be desirable.

The National Ignition Facility. The National Ignition Facility (NIF) will contain the world's most powerful laser. The planned experiments to achieve ignition and high gain for the first time in a laboratory will have far-reaching implications for the future of national security, fusion energy, and a host of scientific and technological fields.

The NIF will focus 192 extremely powerful laser beams onto a pea-sized capsule of deuterium and tritium. Two kinds of targets are under study, both of which can be used in the NIF: *direct-drive targets*, in which a spherical capsule containing deuterium and tritium is struck directly by the laser beams; and *indirect-drive targets*, using a capsule inside a small, thin-walled cylindrical container made from high-atomic-number materials such as gold or lead. The container, called a *hohlraum*, converts the driver beams to x-rays which compress the fuel capsule.

The NIF will also be able to achieve more dramatic results in high-energy-density physics than have been possible with lasers to date. The key physics of certain types of stars, of supernovas, of fluid dynamics and the interior of a nuclear explosion can be revealed, allowing confirmation of computer models and answering questions of nuclear stockpile reliability, fusion energy, and astrophysics.

At a final cost of $2.0 billion, the football stadium–sized facility would be the centerpiece of the nation's inertial confinement fusion research community, leading a worldwide effort to understand the challenging world of high-energy-density physics and possible fusion-energy production. But leading the world into a nuclear testing-free world is also a vital element of the NIF. Its key role in what is described as *science-based stockpile stewardship* responds to a search for alternative means of maintaining confidence in our nuclear deterrent without nuclear testing. The NIF and other elements of the stewardship program allow researchers the insight to help maintain the stockpile in a safe and reliable fashion.

Answers to important energy questions will be possible with the NIF, as one of the key challenges in inertial fusion research is to achieve ignition of the deuterium-tritium fuel capsule, and further, to show a modest gain in the energy produced. Once researchers can show that they produce more energy than the lasers used to create ignition, the inertial confinement community will have vital information on which to base the next step in its fusion energy development.

Groundbreaking research has been under way at Lawrence Livermore National Laboratory using the Nova laser and at the University of Rochester with the Omega upgrade laser, but the boost in technology achieved by NIF (40 times more energy and 10 times the power of nova) will offer information on inertial fusion ignition and gain that is critical to the next era of ICF progress. It will determine the minimum drive energy (and cost) necessary to ignite an inertial fusion target, thus allowing evaluation of the viability of inertial fusion as an energy source. Since the NIF will be able to reproduce conditions that exist in stellar interiors, it will become an important new tool for laboratory astrophysics. It will also allow important experiments in other fundamental sciences.

Initial operation of the first four of 192 beams began in April 2003.

NIF Design Requirements. The design for NIF takes into account the requirements and requests of three main user communities. The top-level performance requirements for the NIF were driven by the indirect- (x-ray-) drive, ICF mission. Those requirements are as follows:

- 1.8 MJ of laser pulse energy on target
- Flexible pulse shaping (dynamic range >50)
- Peak power of 500 TW
- Pulse wavelength in the ultraviolet (0.35 μm)
- Beam power balance better than 8% over 2 ns
- Pointing accuracy $< 50\ \mu$m
- Compatibility with cryogenic and noncryogenic targets
- Ability to do 50 shots per year, each with a yield of 20 MJ, for a total 1200 MJ annual yield
- Maximum credible DT fusion yield of 45 MJ
- Ability to perform classified and unclassified experiments

In addition to these capabilities, weapons physics users desire the highest possible peak power for short pulses ($>$750 TW at 3ω for 1 ns) in order to reach high temperatures and a range of pulse lengths from 0.1 to 20 ns for a wide variety of experiments. These users want bright sources for experiments requiring x-ray backlighting, with small spots at high temperatures (half the energy in a 100-μm spot, and about 95% of the energy at 200 μm). The beams for these backlighters must be pointed a few centimeters away from the center of the main target chamber.

Weapons effects users want the ability to locate arrays of laser targets several tens of centimeters from the target chamber center, as well as 1ω and 2ω capability. Their other requirements include access to the chamber for large, heavy test objects, a well-shielded diagnostics area for testing electronic systems, and no residual light on the test objects.

The NIF target chamber will also have ports that allow the beams to be placed in the proper location for direct-drive ICF experiments and for tetrahedral hohlraums as well as for the baseline cylindrical

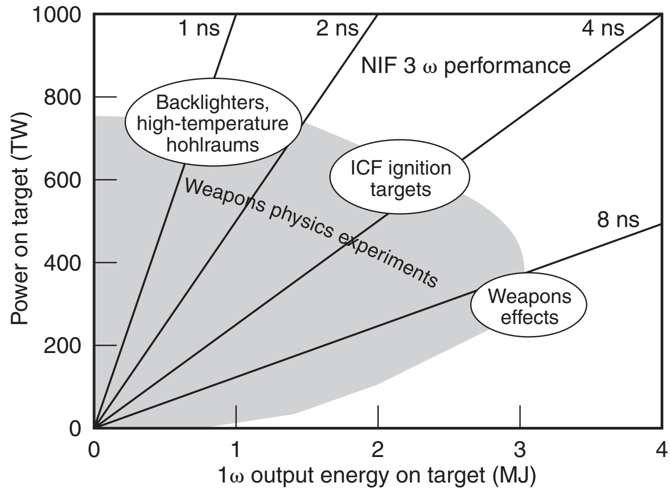


FIGURE 5-56 National Ignition Facility (NIF) users have identified important experiments spanning over a wide range of operating conditions.

hohlraums. All these requirements mean that the NIF must accommodate experiments spanning a wide range of operating conditions (see Fig. 5-56).

NIF Description. The NIF laser system provides routine operation at 1.8 MJ/500 TW in an ignition-target-shaped pulse and has a wide range of operation to meet other user requirements. The laser uses neodymium glass amplifier slabs, with 192 beams in a multipass architecture. Frequency conversion is to the third harmonic: 3ω (350 nm). The laser has adaptive optics (deformable mirrors) to control the beam quality and uses kinoforms and smoothing by spectral dispersion (SSD) to control the beam quality on the target.

Figure 5-57 shows the laser and target area building, as they appear in the Title I Design (submitted to DOE). The overall floor plan is U-shaped, with laser bays forming the legs of the U, and switchyards and the target area forming the connection. The NIF will contain 192 independent laser beams or "beamlets" measuring 40 × 40 cm each. Beamlets are grouped into 2 × 2 "quads," which are stacked two high in 4 × 2 "bundles" (thus eight beamlets per bundle). These bundles are grouped six each into four large "clusters," two in each laser bay, for a total of 192 beamlets (8 beamlets per bundle × 6 bundles per array × 4 clusters). The 192 laser beamlines will require more than 7000 discrete, large optical components (larger than 40 × 40 cm) and several thousand smaller optics. Beams from the laser will strike a series of mirrors, which will redirect them to the large target chamber shown on the right side of Fig. 5-57. The building for the NIF will be about 100 m wide (122 m including the capacitor bays), and about 170 m long.

In deciding how many beams for the NIF, two conditions have to be met. First, there has to be enough beam area facing the target to deliver the required energy. The maximum safe 3ω fluence for an ignition target pulse is about 9 J/cm^2. The maximum practical single beam area is about 1300 cm^2, which would deliver about 11 kJ/beam on the target. At 11 kJ per beam, at least 164 beams are required to put 1.8 MJ on target. Second, the conditions required by indirect- (x-ray-) drive targets are essential factors. These targets (cylindrically symmetric hohlraums) require twice as many beams in the outer cone as in the inner cone, illumination from two directions, and eight or more beam spots per cone. When we multiply these factors together, we find that the beam count must be divisible by 48. The smallest system that satisfies these two conditions is 4 × 48 = 192 beams.

It turns out that it is also very convenient to transport these beams in 48 (2 × 2) clusters, and that this configuration is also compatible with direct-drive uniformity requirements. Finally, 192 beams at 9 J/cm^2 provides 2.2 MJ, a full 20% margin for baseline operation of 1.8 MJ.

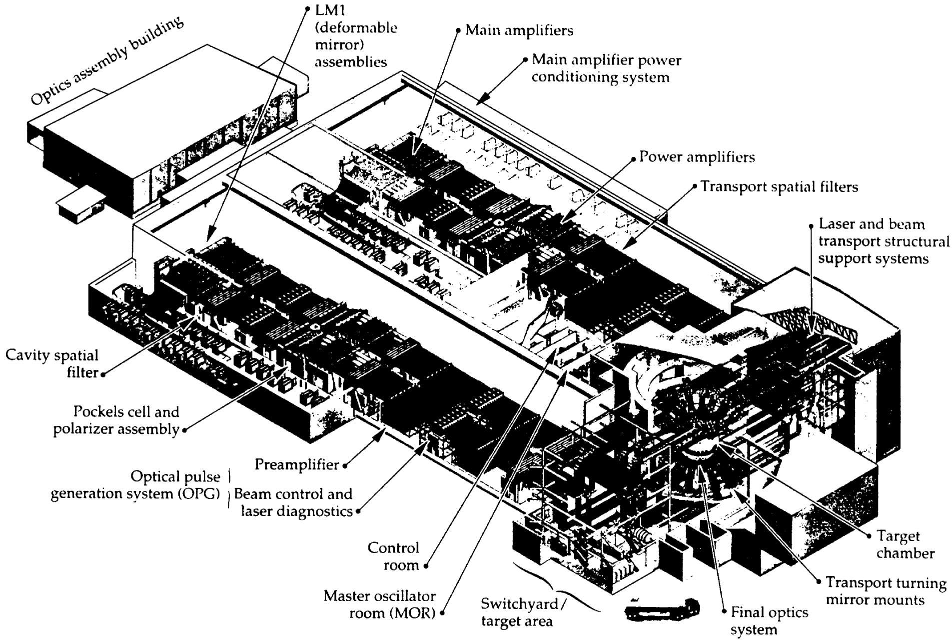


FIGURE 5-57 Layout of the laser and target area building. The two pairs of large beamlet clusters running the length of each laser bay are essentially identical to each other.

Master Oscillator System. The laser pulse is produced in the master oscillator room, where a fiber-ring oscillator generates a weak, single-frequency laser pulse. A phase modulator puts on bandwidth for smoothing by SSD and suppressing stimulated Brillouin scattering. Each pulse is then launched into an optical fiber system that amplifies and splits the pulse into 48 separate fibers. The optical fibers carry the pulses to 48 low-voltage optical modulators very similar to those used in high-bandwidth fiber communication systems. These modulators allow temporally and spectrally shaping of each pulse by computer control. An optical fiber then carries each nanojoule,1-μm pulse to a preamplifier module (PAM).

Preamplifier Module. Optical fibers carrying the pulses from the master oscillator room spread out to 48 preamplifier modules, located on a space frame beneath the transport spatial filters. Each preamplifier has a regenerative amplifier followed by a flashlamp-pumped four-pass rod amplifier. The preamplifier is a two-stage system, designed as a self-contained assembly, that can be pulled out and replaced as needed. The preamplifier brings the pulse to about 10 J, with the spatial intensity profile needed for injection into the main laser cavity. Before entering the main cavity, the output from the preamplifier is split four ways. These four pulses are injected into the four beams that form each of the 48-beam quad arrays.

Main Laser System. Figure 5-58 shows the layout of the main laser components of a NIF beamlet. These components take a laser pulse from the preamplifier to the final optics assembly. A laser pulse from the preamplifier enters the main laser cavity when it reflects from a small mirror labeled LM0. This mirror is located near the focal plane of the transport spatial filter. The 40-cm-diameter pulse exits the transport spatial filter, traveling to the left, and passes through the power amplifier, containing five amplifier slabs.

The beam then enters the periscope assembly, which contains two mirrors (LM2 and LM3) and a switch (a Pockels cell and a polarizer). The pulse reflects off LM3 and the polarizer before passing

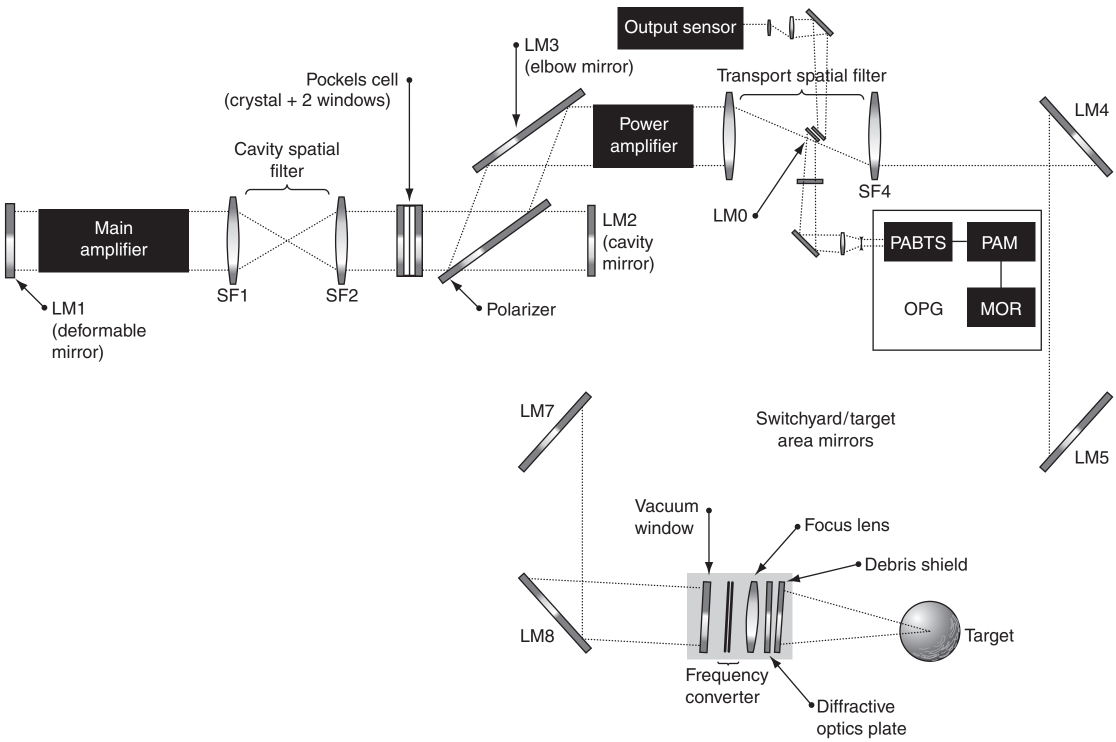


FIGURE 5-58 Schematic diagram of the NIF laser system.

through the Pockels cell. It goes through the cavity spatial filter and the 11-slab main amplifier, and then reflects from a deformable mirror with 39 actuators. After once again passing through the main amplifier, the beam comes back through the cavity spatial filter to the periscope assembly. Meanwhile, the Pockels cell has been fired to rotate the polarization, so the beam passes through the polarizer and strikes mirror LM2, which redirects it back through the cavity spatial filter for another double pass through the main amplifier. The beam returns to the periscope assembly, passes through the deenergized Pockels cell, reflects off the polarizer and LM3, and is further amplified by the power amplifier. Now the pulse passes through the transport spatial filter on a path slightly displaced from the input path. The output pulse just misses the injection mirror LM0 and enters the switchyard and beam transport area.

Switchyard and Beam Transport. Between the switchyards and the target chamber room, the beams are in two 2×2 arrays: the 4×2 bundles are split into 2×2 quads, moving up or down to one of the eight levels of the switchyard. Each pulse now travels through a long beam path, reflecting off of several transport mirrors before reaching the target chamber. The transport mirrors can be moved to create the beam configuration needed for direct- or indirect-drive experiments. For indirect drive, mirrors send the beams straight up or straight down to make the cones coming into the top and the bottom of the target chamber cylinder. For direct drive, 24 beams are directed to circumferential positions around the target chamber by moving two mirrors in each of these beamlines. Once the beams reach the target chamber, they enter the final optics assembly.

Final Optics Assembly. The final optics assemblies are mounted on the target chamber. Each assembly includes a vacuum window at 1ω, a cell that includes a frequency converter (two plates of potassium dihydrogen phosphate crystal) to convert the pulse to 3ω, and the final focusing lens. The cell tips and tilts to tune the frequency converter, and translates along the beam direction to focus the beam on the target. A debris-shield cassette includes the capability of diffractive optics for spot shaping. Once a pulse travels through this assembly, it proceeds to the target in the target chamber.

Other Laser Facilities. Other powerful lasers for ICF experiments include GEKKO XII at Osaka University in Japan, Omega at the University of Rochester in New York, and Nike at the Naval Research Laboratory in Washington, D.C. The GEKKO XII laser used ND: glass laser at 0.53 μm, and 12 beams with random-phase-plates implemented. The laser energy is 8 to 10 kJ with a nominally 1.7-ns (full width at half-maximum, [FWHM]) flat-top pulse with a rise time equivalent to a 1- to 1.3-ns Gaussian pulse (FWHM). Achievements at GEKKO XII include the generation of 10^{13} neutrons per shot, a 10-keV plasma with stagnation-free mode compression, and a density compression up to 600 times solid density using a plastic shell target implosion. An upgrade of the GEKKO system to double the output in blue laser light has been implemented. Further, it is envisioned that a solid-state laser could be ultimately developed for use in a power plant using laser diode (LD) pumping of the laser and advanced active cooling techniques. In addition to technical issues involved in this approach, the projected cost for the LD unit must be substantially reduced to achieve economic competitiveness.

LD-pumped solid-state lasers have advanced to driver candidate status due to the recent advances of AlGaAs LD arrays with power of ~4 kW/cm^2 (200-μs pulse duration), an efficiency of ~60%, and a lifetime of 10^{11} shots. A conceptual power plant design has been carried out for an LD system with 10-MJ blue output, 10% efficiency, and 10-Hz repetition rate. Also, additional constraints incorporated in the design include the following:

To prevent damage, the fluence of the laser beam in the amplifier was restricted to be 32 J/cm^2 at a pulse width of 10 ns.

To prevent parasitic oscillation in the laser disk material, the product of the largest dimension of the disk and the small signal gain coefficient was held $\lesssim 4$.

Cooling was added to keep the maximum temperature difference and the peak temperature in the disk below 10% and 30% of the transition or melting point of the laser material, respectively.

The maximum thermal stress in the disk was held to $\lesssim 20\%$ of the fracture strength to ensure optical quality and long-lifetime operation.

OMEGA Laser Facility. The OMEGA laser is housed in the Laboratory for Laser Energetics (LLE) at the University of Rochester, New York. This facility is part of the national laser fusion effort within the U.S. Department of Energy and has as its main mission direct-drive laser fusion research in support of the upcoming National Ignition Facility (NIF) at the Lawrence Livermore National Laboratory (LLNL). The OMEGA laser system has been operating since May 1994 (McCrory 1994, Seka et al. 1997, Sourers et al. 1996). This 60-beam UV laser facility can cover a wide variety of experiments related to ICF. The facility also supports indirect-drive laser fusion research under the direction of LLNL and the Los Alamos National Laboratory (LANL).

The OMEGA laser system (Boehly et al. 1992) is a Nd:glass laser system with frequency conversion to the third harmonic with a UV (351-nm) energy capability on target in excess of 30 kJ for pulse durations of ≤2 ns. The system's 60 beams (30-cm beam diameter) are arrayed symmetrically around the 3.3-m-diameter target chamber and are focused onto the target by 60 near-diffraction-limited *f*/6 lenses.

Laser pulses with a wide variety of predetermined, temporal UV output pulse shapes have been produced by a versatile pulse-shaping system. These shapes range from 1- to 3-ns flat-topped pulses to linear ramps (1 to 3 ns), Gaussians (0.2 to 1.2 ns), and a special 2-ns pulse designed for indirect-drive targets.

The demonstrated energy balance of the OMEGA system is ~2% rms for the IR part of the system, ~3% rms in the UV after the conversion crystals, and ~4% to 5% rms on target after the UV transport optics (two turning mirrors, a focusing lens, and a blast shield per beamline). The system's reproducibility is <1% rms excluding clearly identifiable flashlamp or PFN malfunctions. The measurement accuracy of the beam energies (including residual fundamental and second-harmonic output from the conversion crystals) is ~0.8% rms.

The OMEGA laser system has three complete driver lines with corresponding pulse-shaping systems. Any one or any combination of two of these driver lines can be used to drive the laser system. The laser system supports the propagation of two different pulse shapes: one for the main experiment (driver) and the other for a backlighter. In this case, one driver line feeds 40 beamlines, while the other feeds the remaining 20. The timing between these two groups of beams can be chosen freely.

Beam smoothing, an essential component for the success of direct-drive laser fusion, is implemented on OMEGA using smoothing by SSD pioneered at LLE. LLE's approach combines continuous-phase-plate technology with polarization rotators and two FM modulators with corresponding spectral dispersion in two dimensions (2-D SSD). Present capabilities are limited to 2- to 3-Å total bandwidth for the two modulators (3 and 3.3 GHz modulation frequency). In addition, a limited number of phase plates and polarization rotators are available primarily for planar target experiments. A full complement of phase plates is available for spherical irradiation experiments, and complete complement of polarization rotators is expected soon. At that time, increased bandwidth (~10 Å) and higher FM modulation frequencies (~10 GHz) will also be available, allowing an on-target frequency bandwidth of ~1 THz in the UV spectrum.

The pointing capability of the OMEGA laser system has been shown to be ~10 μm rms on target. This performance has been repeatedly verified with x-ray imaging using pinhole cameras and x-ray microscopes. The stability of the system is sufficient to maintain the above pointing accuracy without adjustments for at least a period of 1 day. Furthermore, beams can be placed in basically arbitrary locations within ~1 cm from the target chamber center with essentially the same accuracy (~20 mm without corrections; better accuracy is achievable with one iterative fine correction). This capability has been demonstrated on recent indirect-drive hohlraums of cylindrical (two laser entrance holes) and spherical shape ("tetrahedral" hohlraums with four laser entrance holes) with or without additional backlighter targets.

The present OMEGA experiments address irradiation uniformity, implosion physics, hydrodynamic instabilities, laser imprinting, and laser-plasma-interaction physics. In addition, there are experiments carried out by LLNL and LANL relating to indirect-drive laser fusion.

Liquid or solid layers of DT are required to minimize the drive requirements for high-performance, direct- and indirect-drive targets. In a collaboration involving LLNL, LANL, General Atomics (GA), and LLE, the technology required to conduct cryogenic-fuel-layer ICF experiments is being developed at the OMEGA facility (Sourers et al. 1996).

As part of the national effort on cryogenic target development, GA is currently designing a cryogenic target delivery system for OMEGA experiments. The system is shown schematically in Fig. 5-59.

Using this system, polymer shells will be filled with DT up to 1500-atm-pressure, room-temperature equivalent. The shells will be transferred to the vicinity of the OMEGA target chamber in a cold-transfer cryostat. Fuel layers up to 100 μm thick will be formed inside the 1-mm-diameter plastic shells. The capsules will then be inserted into the OMEGA chamber and irradiated by the 60-beam system. Liquid layering is expected to occur inside the chamber with in situ characterization just before the target is shot.

A $50 million upgrade to the OMEGA laser is planned for completion in 2007. It will add two high-energy Petawatt lasers for advanced fast-ignition experiments.

The NIKE KrF Laser. In order to achieve an attractive laser-driver for inertial fusion energy (IFE), a higher laser efficiency is needed than achieved with current flash-lamp-pumped Nd-glan lasers. One approach that combines a reasonably high laser efficiency (around 5%) and short wavelength (UV) is the KrF laser (Sethian et al. 1997).

NIKE is a multikilojoule krypton-fluoride (KrF) laser that has been built at the Naval Research Laboratory (NRL), Washington, D.C. to study the physics of direct-drive inertial confinement fusion. Both the two final amplifiers of the NIKE laser are electron-beam-pumped systems. This design places strong reliance on the pulsed electron beams to drive the amplifiers. The smaller of the two has a 20 × 20-cm aperture and produces an output laser-beam energy in excess of 100 J (Fig. 5-60).

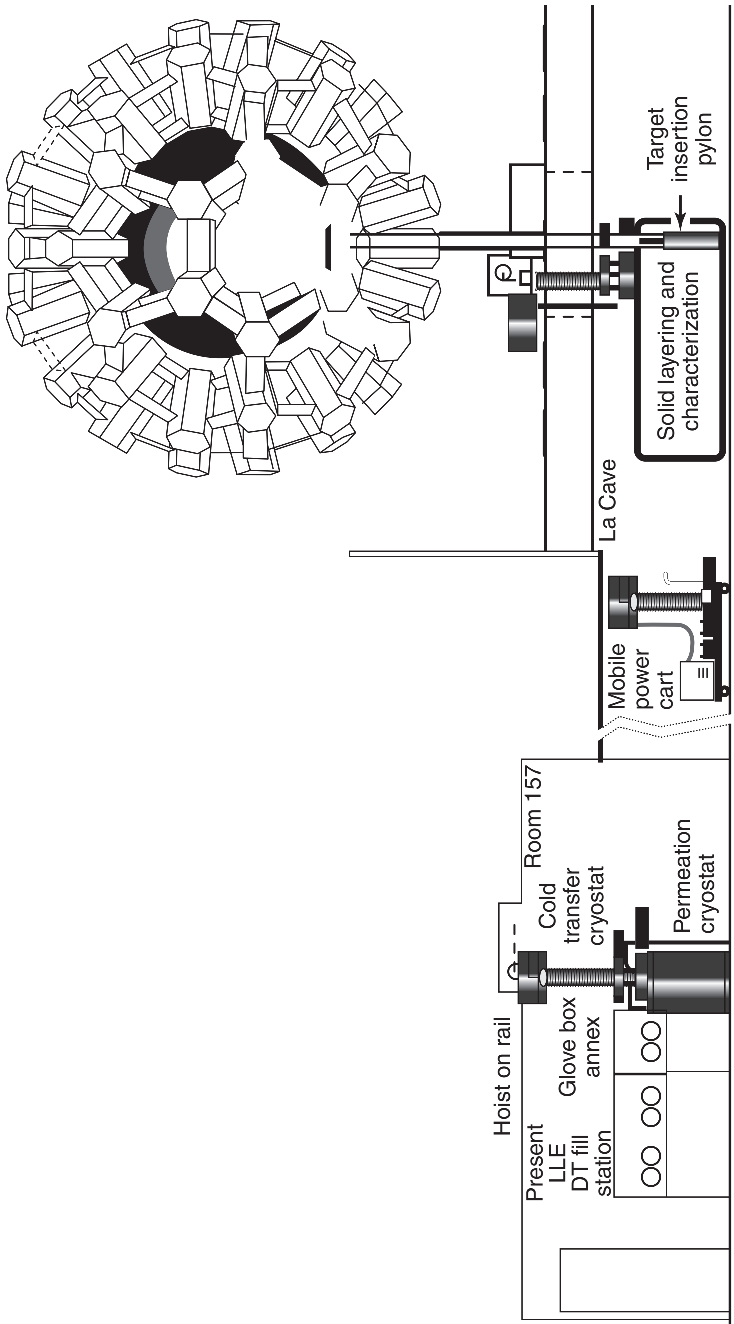


FIGURE 5-59 Schematic diagram of OMEGA cryogenic target system.

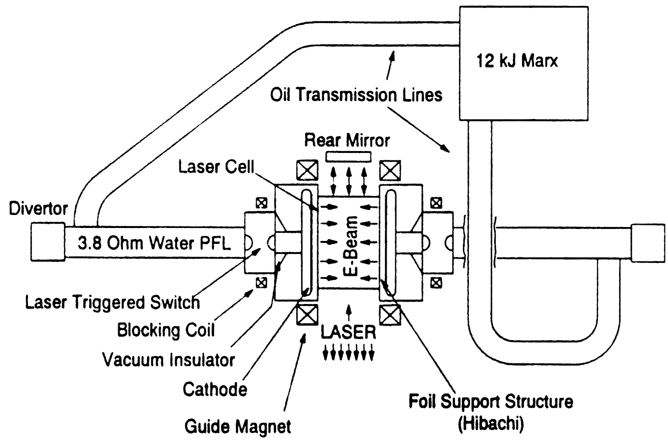


FIGURE 5-60 The NIKE 20-cm amplifier. The paths of the laser beam and electron beams are shown in this figure. The laser beam enters the laser cell (from the bottom of the figure), is amplified once, is reflected off the rear mirror, and is amplified again as it passes back through the cell. The 60-cm amplifier uses a similar concept.

This 20-cm amplifier uses a single 12-kJ Marx generator to inject two 300-kV, 75-kA, 140-ns flat-top electron beams into opposite sides of the laser cell. The larger amplifier in NIKE has a 60 × 60-cm aperture, and amplifies the laser beam up to 5 kJ. This 60-cm amplifier has two independent electron-beam systems. Each system has a 170-kJ Marx generator that produces a 670-kV, 540-kA, 240-ns flat-top electron beam. Both amplifiers are complete, are fully integrated into the laser, meet the NIKE system requirements, and are used routinely for laser-target experiments. To obtain sufficiently high target gains for a direct-drive fusion reactor, a KrF laser based on the NIKE concept must take advantage of three optimizations. First, the laser-beam illumination on the pellet has to be extremely uniform. High-mode beam nonuniformities in the range of 0.2% rms are required, along with low-mode nonuniformities of about 1%. These nonuniformity levels have already been achieved with NRL's KrF laser. Second, the rocket efficiency has to be maximized by depositing the laser energy deeply into the pellet. KrF, with 0.25-μm light, deposits at about twice the plasma density as a solid-state laser at 0.34 μm. Third, the target gain is optimized by "zooming" the laser beam inward during the implosion, thereby better matching the laser spot size to the pellet diameter. This optical zooming is achieved in a straightforward fashion with KrF lasers. There are also several engineering challenges in developing a laser of this type with sufficient energy, repetition rate, reliability, and economy for a practical reactor. Some of these challenges are the lifetime of the emitter and pressure foil in the electron-beam pumped amplifiers, the ability to clear the laser gas between pulses without sacrificing beam quality, and the overall efficiency of the system.

Accelerator and X-Ray Fusion. In addition to lasers, accelerators and x-rays are considered strong candidates for ICF drivers, especially for future reactors. We will consider both briefly here.

Light-Ion Accelerators. Pulsed-power-based light-ion particle accelerators (Bluhm et al. 1992, Mehlhorn 1997, Quintenz et al. 1996) have been under development as ICF drivers since 1972. The technology produces a beam that couples well to matter and is scalable to very high power levels. Light-ion ICF uses pulsed-power accelerators that compress electrical energy in both space and time, and then covert the energy into directed ion beams in a device called an *ion diode*. The ion power and focusability can be improved by using two acceleration gaps, in a device called a *two-stage diode*. Self-pinched transport is an attractive option for propagating the beam from the diode to the target. A stand-off distance of several meters is required to protect the diode from the fusion target blast. Finally, the

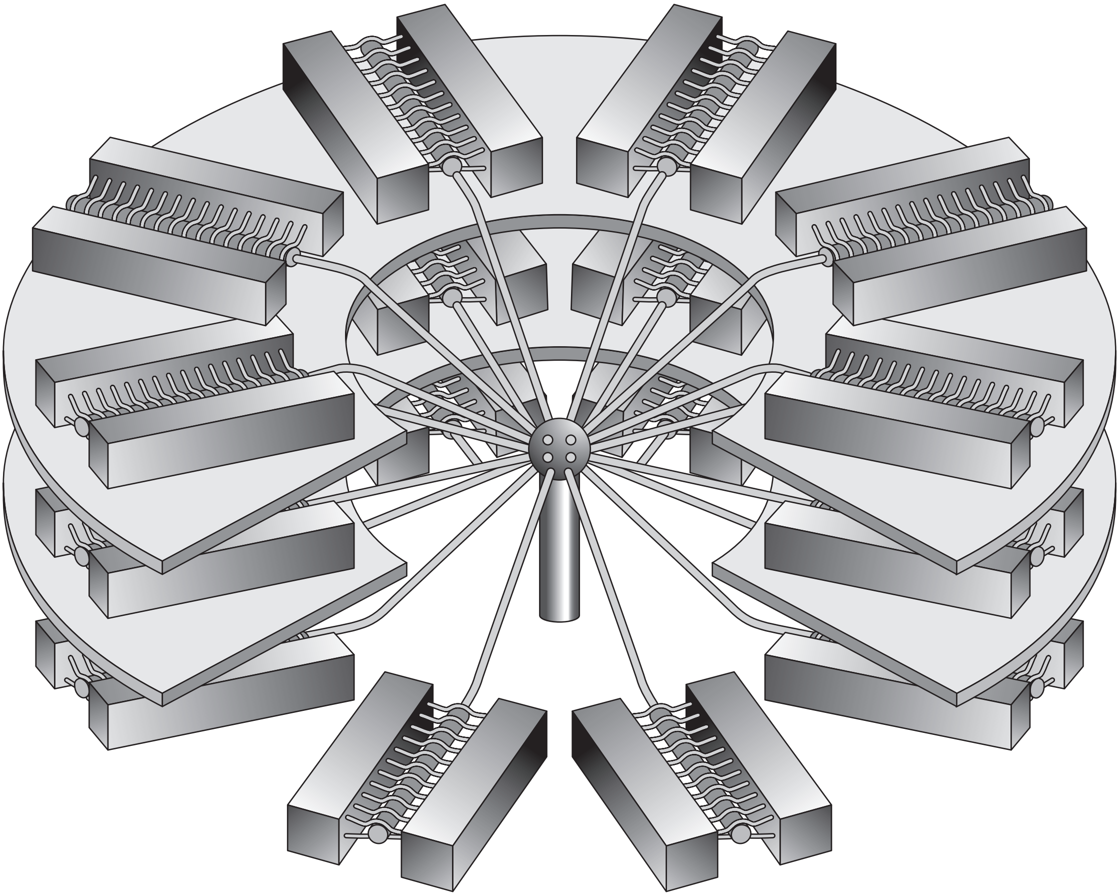


FIGURE 5-61 Z-facility showing modules arranged in three tiers.

light-ion driver beam is focused onto the ICF target, producing x-rays that drive the capsule to fusion conditions, resulting in a large fusion energy yield. Light-ion research is focused on developing the basic physics and technology scalings that could enable the construction of a high-yield facility (HYF) or engineering test facility (ETF) for IFE as shown schematically in Fig. 5-61. A light-ion IFE facility would comprise many modular pulsed-power accelerators that are compact, efficient, and low cost. Ion-beam divergence is the ratio of the transverse to longitudinal beam velocity, and is a measure of the focusability of a beam. Each module would have a two-stage ion diode comprising an injector section that accelerates lithium ions to $\leqslant$24 mrad at 9 MeV, a second stage of acceleration to reduce the divergence to $\leqslant$12 mrad at 35 MeV (by adding only longitudinal momentum), and a self-pinched transport channel that can propagate a 12-mrad lithium beam up to 4 m to the target. The ion beams from these modules would be focused onto a single target to supply the total power and energy requirements for ICF with sufficient uniformity to drive a symmetric implosion. Considerable progress in this technology has been made on accelerators at Sandia National Laboratory (SNL) in New Mexico and the Research Center Karlsruhe (FZK) in Karlsruhe, Germany. Proton beams have been focused to peak intensities of 1 TW/cm^2 at FZK and 5 TW/cm^2 at SNL, while SNL has focused a lithium beam to ~2 TW/cm^2. Focused lithium beams have heated a target to almost 70 eV at specific energy densities that are comparable to the initial pulse level of an HYF target. At present, the key technical challenges for the light-ion program are (1) developing low-emittance ion sources that can provide 1–2 kA/cm^2, (2) reducing total ion beam emittance, (3) increasing total ion beam brightness by two-stage acceleration, and (4) developing a simple robust ion beam transport system. The modularity of the light-ion approach provides a cost-effective development path for an HYF/ETF. The key technical issues can be studied on existing smaller accelerators such as SABRE at SNL and KALIF at FZK; the scaling of the results can be tested on existing larger accelerators such as HERMES-III at SNL, leading to the construction of a single HYF module and the eventual construction of a full HYF/ETF. This technology is now used to produce x-rays for inertial confinement fusion pellet implosions.

Z-Pinch X-Ray Fusion. Since October 1996, the Z-facility (previously known as PBFA-II) at Sandia National Laboratories (Albuquerque) has operated in a Z-pinch configuration to provide experimental data to the Stockpile Stewardship Program of the National Nuclear Security Administration (NNSA). In the new configuration, electrical energy is coupled into the kinetic energy of a magnetically-driven imploding plasma, formed by passing a large current through an annular array of wires. An intense, short pulse of soft x-rays is created when the imploding plasma stagnates on axis. These x-rays are then directed onto the surface of a fusion capsule, heating and compressing it. In a fusion power plant based on Z-pinch x-rays, an array of accelerators would deliver x-rays sequentially to a target chamber at about 0.1 Hz.

Work is proceeding on the design and testing of key systems for the refurbishment of the Z-facility (to be known as ZR) to be complete in 2006.

Heavy-Ion Accelerators for ICF. High-energy particle accelerators are considered as potential ignition devices for inertially driven fusion targets. Relativistic particle beams have the advantage that intrabeam forces are reduced by the cancellation of electric repulsion by magnetic attraction between two particles moving in quasiparallel paths near each other. Also, with large beam energies, huge peak powers can be achieved with relatively small currents. For instance, at 10 GeV, a peak power of 400 TW corresponds to a peak current of 40,000 A (or the combined effect of 40 beams each with 1000-A current), entirely within the realm of achievable currents for a variety of particles.

A key difficulty in the use of high-energy beams is the requirement of depositing most of the beam energy within the relatively thin outer shell of the target. To meet this goal, one can use heavy ions, which even at several GeV have a range smaller than ~1 g/cm^2 or, alternatively, use relativistic electrons, but with an intermediate, "energy-efficient" conversion into soft-x-ray photons (~0.1 to 1.0 keV). The x-rays then drive the target implosion (so-called indirect drive).

The key figure of merit in colliding-beam devices is *luminosity*. This requires high current densities and, in most designs, a large number of high-energy particles packed in a tight bunch. These requirements are not unrealistic, however. For instance, the high current reached many years ago in the CERN-ISR has spurred these ideas on. The energy stored in each beam of the ISR has now reached 5×10^6 J using a beam of 50 A at 31.5 GeV for the duration of 3.3 μs, roughly the correct order of magnitude of energy for ICF ignition. There are, however, two important differences when comparing ICF requirements with CERN-ISR: first, the beam duration for ICF should be $\leqslant$ns in order to achieve the instantaneous power required and second, protons must be replaced by heavier ions (e.g., lithium) in order to ensure a sufficiently short range. These extrapolations are by no means trivial.

To operate a large-scale ICF power station, beams with a total energy delivered to the target of several 10^7 J are required. Since the beam duration must follow the implosion time, the length of the pulse must be a few nanoseconds. Hence, the peak power will reach ~10^{15} to 10^{16} W, or several hundred TW. More detailed calculations with 10-GeV bismuth ions focused to a 2-mm radius target, show that an energy ~5 MJ and a peak power of 300 TW are appropriate for a gain G ~80, sufficient for a commercial-grade reactor.

Another difference between current accelerator experiments and ICF involves the environment in the vicinity of the final focus in ICF. While peak currents in excess of 10^3 A have been focused to submillimetric spots in colliders at hard vacuum without problems, the reactor vessel will have vacuum conditions of 1 to 10 torr. Also, the large amount of radiation emitted by the imploding pellet could "backfire" on the incoming beam.

We can now give a short description of the whole scheme (see Fig. 5-62). The initial acceleration of the singly ionized heavy ions uses a linac structure delivering about 50 mA of current during a

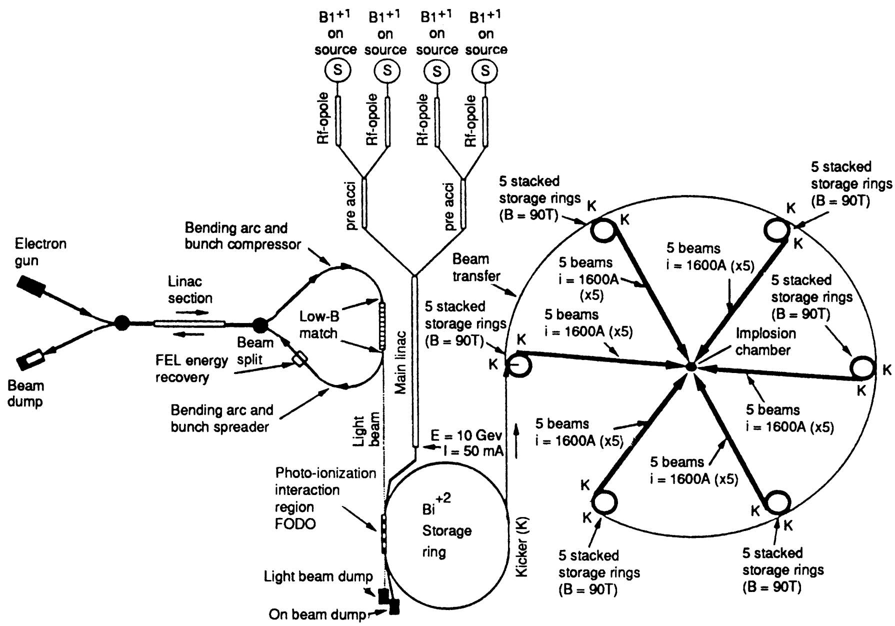


FIGURE 5-62 General layout of the heavy-ion pellet ignition scheme. This is an updated version of the basic HIBALL design.

filling burst of 12.8 ms. To achieve the required current, several sources and preaccelerators are merged. A linac accelerator is complicated and expensive, but compared to other types of accelerators, the single-pass linac appears to be a good choice. The linac pulse is optically dissociated into a doubly ionized beam, which is injected in a storage ring. Every 0.43 ms, the stored beam is dumped into a set of 30 superconducting rings which also act as beam-storage devices. The transfer is performed as a five-turn injection, and the field of these rings is maximized for high compression of the beam. The 30 beams are tightly bunched, extracted through the final focus and the compression channels, and then focused onto the target. At the final phase, each of the 30 beams carries about 1600 A in a pulse duration slightly less than 10 ns.

The U.S. Heavy Ion Fusion Program currently places great emphasis on cost reduction since high capital cost appears to be one of the disadvantages of this approach. The search for lower cost encompasses two main activities: (1) the exploration of new directions in beam physics, plasma physics, and target design (mostly scientific issues), and (2) development of improved materials, better hardware, and better fabrication techniques (mostly technological issues). As the understanding of the science of heavy-ion fusion continues to improve, several developments aid in cost reduction. One involves beam propagation in the target chamber. Recent numerical studies show that beam neutralization techniques will be effective. Beam neutralization allows a reduction in ion kinetic energy, leading to a significant reduction in accelerator cost. Target physics, including fast ignition, appears to be another scientific area where cost breakthroughs are likely. Important cost reductions in most areas of technology also appear feasible. A heavy-ion induction accelerator consists of a number of components: ion sources, an injector, ferromagnetic cores, insulators, magnetic and electric focusing lenses, pulsers, and diagnostics and controls. Although the ion sources by themselves are not costly, improved sources can lead to lower cost in the remainder of the accelerators; consequently, several new source technologies, including aluminosilicate, contact ionization, gas, laser, and spark sources, are being explored. The cost of ferromagnetic materials for the induction cores ranges from a few dollars per kilogram to more than $100/kg. Techniques are under study to insulate and anneal inexpensive materials to make them suitable for accelerator applications. Traditionally, insulators for accelerators are made from alumina.

The joints to metal are made by brazing. Large alumina insulators needed for fusion accelerators are costly, as is brazing. Thus, insulators that can be cast with metal attachments are being developed. Innovative superconducting magnetic lenses with an emphasis on designs that can be inexpensively fabricated are being developed. Conventional and advanced pulsers for heavy-ion fusion are also under study. In terms of conventional pulsers, most work is with thyratrons, exploring methods to optimize them for our application. Advanced pulsers employ solid-state devices and magnetic pulse compression. Finally, improved diagnostics and control systems are being developed. Improved ability to sense and control the beam leads to smaller less expensive accelerators.

Fusion Breeder. The physical basis of the fusion breeder (also called the fusion-fission hybrid) is the prolific production of high-energy neutrons in fusion reactions. These neutrons can be used to breed fertile fuel (plutonium and ^{233}U) for use in conventional fission reactors. Each D-T fusion reaction produces 14 MeV of total energy, or approximately 4 times as many neutrons per unit of energy as a fission event (~3 neutrons per 200 MeV). The 14-MeV fusion neutrons produce additional neutrons in the breeder blanket by neutron-multiplying nuclear reactions, by fast fissions in fertile materials, and/or by nonfissioning reactions such as (n, 2n) reactions in beryllium.

5.4.10 Breeder Types

Two different fusion breeder types based on these two methods of neutron multiplication have emerged. One uses a fast-fission blanket and the other uses a suppressed-fission blanket. In the former, the D-T fusion source is surrounded by a blanket of fertile material (^{238}U and/or ^{232}Th) and a lithium compound for tritium breeding as shown in Fig. 5-63. Fast fissions in the fertile material multiply both the fusion energy (3 to 10 times) and the neutrons (approximately two to four neutrons per fusion neutron). One of the neutrons is needed to breed tritium from lithium and the remainder are available for breeding fissile fuel.

In the suppressed-fission blanket (see Fig. 5-63), a nonfissioning, neutron-multiplying material such as beryllium replaces most of the fertile fuel in the blanket. Consequently, the fast-fission source of neutron multiplication is replaced with a nonfissioning source of neutrons. To further suppress the fissioning of the bred ^{233}U, its concentration is limited to about 1% or less of the fertile fuel. Thus, the breeding blanket must be designed to allow for online, or quick, low-cost removal and recovery of the bred fissile fuel. The low fission rate results in superior overall reactor safety characteristics. Also, a much lower fission product inventory and a lower decay afterheat results.

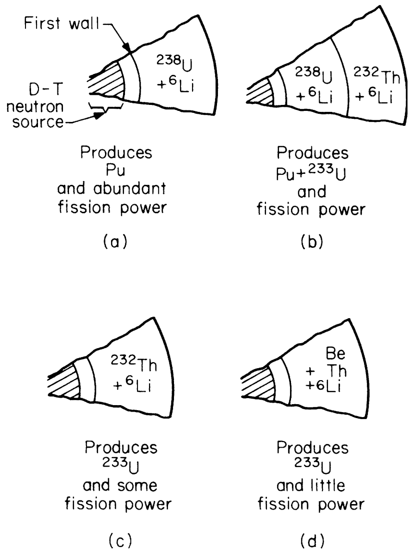


FIGURE 5-63 Methods of neutron multiplication: (*a*) ^{238}U fast-fission blanket; (*b*) ^{238}U + ^{238}Th fast-fission blanket; (*c*) ^{232}Th fast-fission blanket; (*d*) ^{9}Be fission-suppressed blanket.

Fuel Cycles. The suppressed-fission blanket can produce plutonium rather than ^{233}U by substituting uranium carbide for the thorium, but then the number of client LWRs that can be fueled is lower. For example, for a recent design with a 2100-MW fusion driver (4000 MW$_{nuclear}$), the suppressed fission design could support 14 LWRs or 4000 MW$_{nuclear}$ each when producing ^{233}U but only about 9 when producing plutonium. The advantages of higher LWR support motivate the desire for new fuel-reprocessing technologies for ^{233}U and thorium. Two technologies being considered are the conventional, aqueous THOREX reprocessing technology and a pyrochemical process that uses magnesium dissolution of the thorium leaving the ^{233}U as a precipitate.

Breeder Role. Fusion breeders are best understood in the context of a symbiotic fusion-breeder–fission-burner system that generates electricity. The fusion breeder would be incorporated into a fuel-cycle complex along with fuel-reprocessing plants, fuel-fabrication facilities, and possibly a waste-disposal facility—all in a safeguarded area. The fissile fuel produced in the fusion breeder is recovered by reprocessing, mixed with fertile fuel, fabricated into fuel rods, and shipped to the fission burner reactors. The spent fuel from the burner reactors is returned to the safeguarded fuel-cycle complex where the remaining fissile fuel is separated from the radioactive waste material.

Support Ratio. The thermal support ratio is defined as the nuclear (or thermal) power of the client fission reactor (e.g., LWR) divided by the nuclear power of the fusion breeder. A high thermal support ratio is advantageous because such a breeder could fuel a large number of fission reactors, thus having a large commercial impact. Also, for large support ratios, only a small fraction of the symbiotic system's cost and electricity generation is attributed to the addition of a fusion breeder into the existing electricity generation system (typically 15% of the overall capital cost and about 5% of the overall electricity generation). Because relatively little power would be produced within the safeguarded fuel-cycle park itself, the breeders could be in a remote location.

Thermal support ratios for fusion breeders range from 4 to 45, depending on the choice of fusion blanket and client thermal converter (e.g., LWRs or advanced converter reactors). The following support ratio estimates are typical: uranium fast-fission blankets produce enough plutonium to support 4 to 6 LWRs; thorium fast-fission blankets produce enough ^{233}U to support 8 to 12 LWRs, or 14 to 28 advanced converters; uranium suppressed-fission blankets produce enough plutonium to support 9 to 11 LWRs; thorium suppressed-fission blankets produce enough ^{233}U to support 12 to 16 LWRs, or 35 to 45 advanced converters.

The variations in these support-ratio estimates depend on the specifics of the fusion-blanket designs, the type of client fission reactor, and fuel-cycle choices.

These support ratios can be put in perspective by comparing them to the support ratios of a liquid-metal fast-breeder reactor (LMFBR). A typical LMFBR does not produce enough excess fissile fuel to support even one LWR of equivalent nuclear power. Furthermore, the LMFBR must also produce fissile fuel to satisfy the fissile inventory requirement of additional LMFBRs. Consequently, LWR support is not an effective mode of LMFBR operation. The fusion breeder, on the other hand, requires no initial fissile inventory, and its tritium inventory is quite low.

Performance Requirements. Fusion breeders can operate economically with significantly lower fusion performance and higher cost than fusion electric power plants. We consider two performance indicators for magnetic fusion devices: plasma energy gain Q and first-wall fusion neutron wall loading Γ. Q is defined as the ratio of the fusion energy produced to the input energy required to heat and sustain the plasma (supplied by relatively expensive beams of energetic neutral atoms and/or rf heating systems). The fusion neutron first-wall loading in megawatts per square meter is indicative of the blanket power density.

Several conceptual design studies of fusion reactors have shown that pure fusion power plants will require $Q > 20$ and $\Gamma > 3$ MW/m^2 to produce competitive electricity. Although high Q is preferred, fusion breeders will only require Q values from 2 to 6 and $\Gamma > 1$ MW/m^2 to economically produce fuel for LWRs. Fusion breeders with fast-fission blankets can operate in the lower-Q regime (i.e., Q from 2 to 4); with suppressed-fission blankets, Q values above 6 are required.

These performance requirements also apply to ICF systems. For example, ICF hybrid studies have shown that a driver-efficiency target gain product ($\eta_D G$) of about 6 is acceptable while an $\eta_D G$ product greater than 20 is required for fusion electric power generation. These lower performance requirements can allow relaxation of technology requirements.

5.4.11 Progress toward Attainment of Controlled Fusion

Figure 5-64 illustrates the progress of magnetic fusion toward the goal of controlled nuclear fusion. In this figure, the kinetic temperature of a pure deuterium-tritium plasma (ions and electrons taken to have the same temperature) is plotted against the density-confinement time product $n\tau$. In these coordinates, ignition occurs at any point on or above the locus (top of the diagram) at which the thermonuclear energy gain exceeds 30.

Loci of constant gain (fusion power density at 17.6 MeV per fusion, divided by the external power density required to maintain the plasma temperature against losses) are plotted over a wide range (from

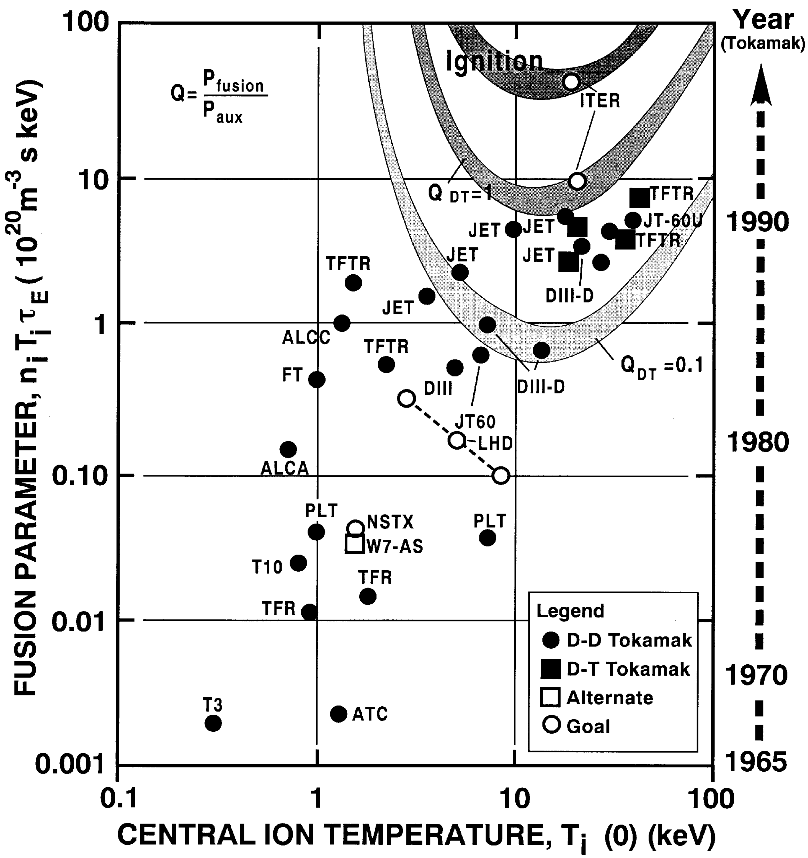


FIGURE 5-64 Progress in magnetic fusion. [***Legend:*** ALCATOR (ALCA and ALCC) = high-field tokamak, MIT; ATC = adiabatic toroidal compressor tokamak, Princeton; DIII = doublet tokamak, GA Technologies, Inc.; JET = Joint European Torus, Culham, U.K.; NSTX = National Spherical Tokamak Experiment, Princeton; START = Spherical Tokamak, Culham, U.K., Lawrence Livermore Laboratory; ORMAK = Oak Ridge National Laboratory tokamak; PLT = Princeton Plasma Physics Laboratory, large-torus tokamak; TFR = tokamak Fontenay-aux-Roses, France; TFTR = tokamak fusion test reactor, Princeton; ITER = international thermonuclear experimental reactor (proposed international project—status is uncertain due to global financial issues); 22 = IIB-mirror reactor, Lawrence Livermore National Laboratory.]

10^{-8} to 10) over which fusion confinement experiments have operated, or are projected to operate. The values of the operating areas of specific experiments are identified by the abbreviated symbols for several confinement devices. The highest attained kinetic temperature achieved is in mirror and certain pinch experiments where about 30 keV, equivalent to 300 million degrees absolute, has been achieved.

Unique Resource Requirements. Table 5-19 lists a number of materials needed to support a nuclear-fusion power economy capable of delivering a total power of a million megawatts (third column) compared with the forecast U.S. demand for these materials in the year 2000 and estimates of world reserves. According to this table, there appears to be no resource limitation that would prevent the extensive development of fusion power. However, the demands for certain materials (beryllium, copper, helium, lithium, vanadium, niobium, and molybdenum) would require additional production capacity, and exploration for new sources would be required for beryllium, lithium, vanadium, niobium, and molybdenum.

TABLE 5-19 Estimated Quantities of the Unique Resource Requirements Associated with Fusion Power

Material	Application	Inventory for 10^6 MW capacity, metric megatons	Forecast U.S demand for year 2000, metric megatons	World reserves, metric megatons	Quantity contained in upper 10 m of Earth's crust metric megatons	Ratio of quantity in Earth's crust to world reserves
Beryllium	Neutron multiplication and PFCs*	0.046	0.002	0.09	1.5(4)†	1.6(5)
Aluminum	Coil conductor	1.0–2.6	33	1000	5.1(8)	5.1(5)
Copper	Coil and firewall conductor	3.2–8.6	7	300	2.3(5)	7.6(2)
Helium	Refrigerant	0.04–1.1	0.02	1	4(3)‡	4(3)
Lithium	Fuel, coolant	0.95–1.5	0.014	0.8	2(5)§	2.5(5)
Titanium	Structure, SC¶	0.5	2	150	2.8(7)	1.9(5)
Vanadium	Structure, SC	2.4	0.03	10	6.6(5)	6.6(4)
Niobium	Structure, SC	3.3	0.01	10	9.4(4)	9.4(3)
Molybdenum	Structure	2.8	0.09	5.4	6.2(3)	1.1(3)
Tin	SC	0.3	0.1	7.2	1.4(4)	2(3)
Lead	Shielding	11	2.5	95	9.4(4)	1(3)

*PFCs = Plasma-facing components.
†To be read as 1.5×10^4.
‡In the atmosphere.
§There is approximately the same quantity of lithium in seawater.
¶SC = Superconductive materials.

Many of these materials will require new fabrication technology and the development of measures against structural deterioration under neutron bombardment and other radiation effects.

Safety and Environmental Considerations. Fusion power has long been considered preferable to fission power from the viewpoints of safety and environmental impact. Fusion reactors, in large part because they are concerned with fuels and reaction products of low atomic number, have lower potential for the release of radioactive materials. Also, the fusion process does not inherently result in radioactive products that require long-time storage like the fission-product waste from fission reactors. But more detailed examination of the processes involved in fusion shows that there are several key points at which stringent radiation safety measures must be taken.

The principal environmental hazard of fusion reactors employing the D-T-Li cycle is the release of tritium during normal operation or in the event of an accident. The avenues of escape of tritium during normal operations are the lithium blanket and its associated containment structure, fluid piping, and heat-exchanger tubing. All these elements operate at extremely high temperatures, with corresponding high permeation rates.

Fusion safety standards have been developed that require tritium releases to be maintained below 10 Ci/day in routine operations and considerably higher values are proposed in the event of a severe accident. It is assumed, on the basis of experimental evidence, that essentially all elemental tritium escaping to the environment will become oxidized. Oxides are several thousand times more hazardous than the element itself.

In a 1000-MW fusion plant, the internal throughput of tritium of about 4×10^6 Ci/day would require a containment factor of 99.9999% to limit the escape to 10 Ci/day. While such a factor is feasible, its attainment at acceptable costs must be demonstrated before large-scale fusion plants employing tritium can be demonstrated to be cost effective.

Fusion neutrons will also cause activation of the first-wall and blanket materials. This radioactivity can be minimized by judicious selection of materials. The overall biological-hazard potential of a typical fusion blanket system is lower than that of a liquid-metal fast-breeder fission reactor by at least one order of magnitude. While this fact does not in itself demonstrate that fusion reactors are safer than fission reactors, it does indicate that the technical problems of ensuring fusion safety probably will be simpler. Both types of reactor require stringent design constraints in this respect.

The problems of an accidental or forced shutdown involve two major factors: (1) the afterheat associated with loss of coolant of the plasma-facing structure and in the lithium blanket and its structural elements, and (2) the energy storage of the magnets in the confinement system. The afterheat depends markedly on the structural materials (silicon carbide, vanadium, and ferrite steels are the three current candidates) associated with the blanket. The duration of the cooling period may vary from hours to years. These materials have biological-hazard potentials lower than plutonium, but still they are not negligible.

The forces associated with the large magnet structures at the high fields required for fusion are of large proportions. During normal opration, for example, in a typical tokamak reactor, the force tending to push the toroidal field coils toward the center is 10^8 lb per coil at a field of 150 kG.

The stored energy in the system can be as high as 10^{11} J. Uncontrolled quenching of the field can produce extremely high voltages, so the insulators associated with the magnetic system must have a high safety factor. Even in normal operation, when the field energy is purposely dissipated by passage through a resistance bank, a residual amount of the energy, of the order of 1%, may remain. This residue (10^9 J of energy) is still enormous by conventional standards.

The foregoing examples illustrate the nature and extent of the engineering problems that will remain to be solved after controlled fusion itself is demonstrated and reduced to practice.

BIBLIOGRAPHY

Bell, M. G., et al., 1997. Deuterium-Tritium Plasmas in Novel Regimes in the Tokamak Fusion Test Reactor, *Phys. Plasmas*, May, vol. 4.

Bertolini, E., 1995. JET with a Pumped Divertor: Design, Construction, Commissioning and First Operation, *Fusion Eng. Design*, vol. 30, pp. 53–66.

Bertolini, E., and The JET Team, 1997. Current Engineering Issues and Further Upgrading of The JET Tokamak, *Proc. 17th Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE, p. 81.

Bluhm, H., Hoppe, P., Althaus, M., Bachmann, H., Bauer, W., Baumung, K., Buth, L., Karow, H. U., Laqua, J., Rusch, D., Stein, E., and Stoltz, O., 1992. Focusing Properties of a Strongly Insulated Applied-Br-Proton Diode with a Preformed Anode Plasma Source, *Proc. 9th Int. Conf. High-Power Particle Beams*, pp. 51–67.

Boehly, T. R., et al., 1992. In Quarles, G. J., ed., *Solid State Lasers III*, Bellingham, Wash., SPIE, vol. 1627, pp. 236–245.

Boozer, A., Lyon, T., and Skebet, L., 1998. Role of Stellarators in the U.S. Fusion Program. Report for the DOE SCICOMIAC-Panel (see http://www.aries.ucsd.edu/SCICOM/AC-Panel).

Calis, R., et al., 1989. DIII-D Results and Implications, *Fusion Technol.*, vol. 15, no. 2, pt. 2, pp. 275–278.

Campbell, E. M., Hogan, W. J., and Lowdermilk, W. H., 1992. Nova Upgrade Mission and Design, *Fusion Technol.*, vol. 21, no. 3, pt. 2A, pp. 1344–1349.

Carpignano, F., Coppi, B., Nassi, M., and the Ignitor Project Group, 1995. The Ignitor Machine, *Fusion Technol.*, Herschbach, K., Mauere, W., and Vetter, T., eds., New York, Elsevier Press.

Choi, C. K., Gilligan, J. G., and Miley, G. H., 1979. *The SAFFIRE D-^{3}He Pilot Plant Concept.* Annual Report, EPRI Research Project 645-1, Palo Alto, Calif., EPRI.

Conn, R. W., 1983. The Engineering of Magnetic Fusion Reactors, *Sci. Am.*, vol. 249, no. 4, pp. 60–71.

Conn, R. W., Chuyanov, V. A., Inoue, N., and Dr. Sweetman, R., 1992. The International Thermonuclear Experimental Reactor, *Sci. Am.*, April, vol. 266.

Cook, D. L., 1992. Results from PBFA II, *Fusion Technol.*, vol. 21, no. 3, part 2A, p. 1358.

Coppi, B., Nassi, M., and Sugiyama, L. E., 1992. *Phys. Scripta*, vol. 45, no. 112.

Crandall, D. H., 1992. Conclusions and Directions from the OFE Inertial Fusion Reactor Studies, *Fusion Technol.*, vol. 21, part 2A, p. 1451.

Dean, S. O. 2002, Fifty Year of Fusion Research, *Nuclear News*, July 2002, pp. 34–41.

Dean, S. O., Frontiers in Fusion Research, *J. Fusion Energy*, vol. 19, nos. 3/4, pp. 293–302.

Dolan, T. J., 1982. *Fusion Research*, New York, Pergamon Press.

Erckman, V., et al., 1996. The W7-X Project: Scientific Basis and Technical Realization, *Proc. 17th IEEE/NPS Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE.

Fowler, T. K., Hardwick, J. S., and Jarboe, T. R., 1994. *Comments Plasma Phys. Contr. Fusion*, vol. 16, p. 91.

Goldston, R. J., A plan for the Development of Fusion Energy, *J. Fusion Energy*, vol. 21, no. 2, pp. 61–112.

Gu, Y., and Miley, G. H., 1995. Spherical IEC Device as a Tunable X-Ray Source, *Bull. APS*, vol. 11, p. 1851.

Hagenson, R. L., and Krakowski, R. A., 1985. *Fusion Technol.*, vol. 8, p. 1606.

Harkness, S. D., Review of the Fusion Materials Research Program, *J Fusion Energy*, vol. 19, no. 1, pp. 45–64.

Haruhiko, H., et al., 1995. Rethermalization of a Field-Reversed Configuration Plasma in Translation Experiments, *Phys. Plasmas*, vol. 2, p. 191.

Hemmerich, J. L., et al., 1989. Key Components of the JET Active Gas Handling System: Experimental Program and Test Results, *Fusion Technol.*, vol. 1, pp. 93–100.

Hoffman, A. L., 1996. An Ideal Compact Reactor Based on a Field Reversed Configuration, *Fusion Technol.*, vol. 30, p. 1367.

Hoffman, A. L., et al., 1993. The Large S Field-Reversed Configuration Experiment, *Fusion Technol.*, vol. 23, p. 185.

Hogan, W. J., ed., 1995. *Energy from Inertial Fusion,* Vienna, Austria, IAEA.

Hooper, E. B., Innovative Confinement Concepts Workshop—2002 Conference Report, *J. Fusion Energy*, vol. 21, nos. 1/2, pp. 13–18.

Hooper, E. B., and Fowler, T. K., 1996. Spheromak Reactor: Physics Opportunities and Issues, *Fusion Technol.*, vol. 30, pp. 1390–1394.

Huguet, M., et al., 1987. The JET Machine: Design, Construction and Operation of Major Subsystems, *Fusion Technol.*, vol. 11, p. 43.

IEEE, 1997. *Proc. 17th Symp. on Fusion Engineering*, San Diego, Calif.–New York, IEEE.

Irby, I., and The Alcator Group, 1997. The Alcator C-Mod Tokamak and Recent Results, *Proc. 17th IEEE/NPSS Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE.

ITER, 1998. *Conceptual Design Activities Final Report*, IAEA/ITER/DS/16, Vienna, Austria, IAEA.

ITER (U.S.) Home Team Group, 1998. *ITER Final Design Report*, IC-33, Project Office, University of California, San Diego.

Johnson, D., et al., 1995. Recent D-T Results on TFTR, 22d European Physical Society Conference, Bournemouth, U.K., July.

Keilhacker, M., and the JET Team, 1997. JET D-T Experiments and Their Implications for ITER, Current Engineering Issues and Further Upgrading of the JET Tokamak, *Proc. 17th Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE.

Kulcinski, G. L., et al., 1997. Discharge Characteristics of the Spherical Inertial Electrostatic Confinement (IEC) Device, *IEEE Transactions on Plasma Science*, vol. 15, no. 4, pp. 733–739.

LaMarche, P. H., et al., 1994. Tritium Processing and Management during D-T Experiments on TFTR, *Fusion Technol.*, vol. 26, no. 3, Pt. 2, pp. 427–433.

Lawrence Livermore National Laboratory, National Ignition Facility, 1997. URCL-LR-105821-97-3, ICF Quarterly Report, Livermore, Calif., LLNL.

Logan, B. G., 2003, Status of Heavy Ion Fusion Research, posted at http://www.fire.pppl.gov/fpa_annual03.html.

Logan, B. G., and Meier, W. R., 1998. Impact of Fast Ignition for Inertial Fusion Energy, Abstracts, *13th Am. Nucl. Soc. Topical Mtg. on Tech. of Fusion Energy*, Am. Nuc. Soc., June, p. 14.

Martin, G., 1997. Enhanced Performances for Long Pulses in Tore Supra, *Proc. 17th IEEE/NPSS Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE.

McCarthy, K., et al., Non-electric Applications of Fusion, *J. Fusion Energy*, vol. 21, no. 3/4, pp. 121–154.

McCrory, R. L., et al., 1996. Direct Drive Laser Fusion Experimental Program at the University of Rochester's Laboratory for Laser Energetics, *Plasma Physics and Controlled Nuclear Fusion Research*, Vienna, Austria, IAEA, vol. 3, pp. 33–37.

McCrory, R. L., 2003, Status and Plans for Omega 2003, posted at http://www.fire.pppl.gov/fpa_annual03.html.

Mehlhorn, T. A., 1997. Intense Ion Beams for Inertial Confinement Fusion, *IEEE Trans. Plasma Sci.,* vol. 25, pp. 1336–1356.

Miley, G. H., 1980. Overview of Nonelectrical Applications of Fusion, *Proc. 2d Miami Int. Conf. Alternate Energy Source*, vol. 6, no. 19, p. 2585.

Miley, G. H., 1997. The Inertial Electrostatic Confinement Approach to Fusion Power, in Panarella, E., ed., *Trends in International Fusion Research*, New York, Plenum Press, pp. 135–148.

Miley, G. H., and Campbell, M. E., eds., 1997. Laser Interaction and Related Plasma Phenomena, *AIP Conf. Proceedings 406*, Woodbury, N.Y., American Institute of Physics.

Moir, R., 1992. HYLIFE-II Inertial Fusion Energy Power Plant Design, *Fusion Technol.*, vol. 21, p. 1475.

Moir, R. W., 1996. *Liquid First Walls for Magnetic Fusion Energy*, LLNL Report UCRL-ID-123902, March 28.

Momota, H. G., et al., 1992. Conceptual Design of the D-^{3}He Reactor Artemis. *Fusion Technol.*, vol. 21, pp. 2307–2323.

Motojima, O., et al., 1993. *Fusion Eng. Design*, vol. 20, no. 3.

Motojima, O., et al., 1995. Long Pulse Operations Scenario of LHD, *Trans. Fusion Technol.*, vol. 27, no. 123.

Najmabadi, F., and the ARIES team, 1997. Overview of the ARIES-RS Reversed-Shear Tokamak Power Plant Study, *Fusion Eng. Design*, October.

Nakai, S., Mima, K., and Kitagawa, Y., 1992. Status and Plans for GEKKO XII, *Future Technol.*, vol. 21, p. 135.

Nazikian, R., et al., 1996. Observation of Alpha-Particle Driven Toroidal Alfven Eigenmodes in TFTR D-T Plasmas, *16th IAEA Fusion Energy Conf.*, Montreal, Canada, October.

Neilson, G. H., The Fusion Science Research Plan for the Major U.S. Tokamaks, vol. 18, no. 3, pp. 117–160.

Neyatani, Y., and The JT-60 Team, 1995. Recent Progress in JT-60U Experiments, *Fusion Eng. Design*, vol. 30, pp. 25–37.

Olson, C., 2003, Status of Z-Pinch Fusion posted at http://www.fire.pppl.gov/fpa_annual03.html.

Peng, M., Status of Spherical Torus Research, posted at http://www.fire.pppl.gov/fpa_annual03.html.

Perry, M. D., et al., 1996. *CLEO'96,* Technical Digest Series, Washington, D.C., Optical Society of America, vol. 9, p. 307.

Petersen, P. I., and The DIII-D Team, 1997. Recent Results from the DIII-D Tokamak, *Proc. 17th IEEE/NPSS Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE.

Quintenz, J. P., Bloomquist, D. D., Leeper, R. J., Mehlhorn, T. A., Olson, C. L., Olson, R. E., Peterson, R. R., Matzen, M. K., and Cook, D. L., 1996. Light Ion Driven Inertial Confinement Fusion, *Prog. Nucl. Eng.*, vol. 30, pp. 183–242.

Robins, J., et al., 1993. Tritium Purification System for TFTR, *Proc. 15th IEEE/NPSS Symp. on Fusion Engineering,* Hyannis, Mass., October.

Sakasai, A., and The JT-60 Team, 1997. High Performance and Steady-State Experiments on JT-60U, *Proc. 17th IEEE/NPSS Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE, pp. 69–72.

Seka, W., et al., 1997. *OMEGA Experimental Program and Recent Results. Laser Interaction and Related Plasma Phenomena*, 13th Int. Conf., In Miley, G. H., and Campbell, E. M., eds., American Institute of Physics, pp. 56–65.

Sethian, J. and Obenschain, S., Fusion Energy with Lasers, Direct Drive Targets, and Solid Wall Chamber, posed at http://www.fire.pppl.gov/fpa_annual03.html.

Sethian, J. D., et al., 1997. The NIKE Electron-Beam-Pumped KrF Laser Amplifiers, *IEEE Trans. Plasma Sci.*, vol. 25, pp. 221–227.

Sourers, T. M., et al., 1996. The Role of LLE in the NIF Project, *Fusion Technol.*, vol. 30, no. 3, pt. 2A, p. 492.

Stork, D., and The JET Team, 1997. JET Engineering Development towards D-T Operations in an ITER-Like Machine Configuration, *Proc. 17th IEEE/NPSS Symp. on Fusion Engineering*, San Diego, Calif.–Piscataway, N.J., IEEE.

Spears, W. R., 2003, Status of ITER posted at http://www.fire.pppl.gov.fpa_annual03.html.

Sved, J., 1997. The Commercial IEC Portable Neutron Source, *Trans. of the ANS*, vol. 77, p. 504.

Tabak, M., et al., 1994. *Phys. Plasmas*, vol. 1, no. 5, p. 1626.

The JET Team, 1992. Fusion Energy Production from Deuterium-Tritium Plasma in The JET Tokamak, *Nucl. Fusion*, vol. 32, p. 187.

The JET Team, 1995. The New Experimental Phase of JET and Prospects for Future Operation, *Plasma Phys. Controlled Nucl. Fusion Research*, vol. 1, pp. 51–79.

Tolok, V. T., A History of Stellarators in the Ukraine, *J. Fusion Energy*, vol. 20, no. 4, pp. 117–130.

Von Halle, A., 1998. Final Operations of the Tokamak Fusion Test Reactor (TFTR), *Proc. 17th IEEE Symp. on Fusion Engineering*, San Diego, Calif.–New York, IEEE.

Wagner, F., 1998. Stellarators and Optimized Stellerators, Trans. *Fusion Technol.*, vol. 33, pp. 67–83.

Wakatani, M., 1998. *Stellarator and Heliotron Devices*, Oxford University Press.

Warner, B. E., 2003, NIF: Transition to Target Shooter posted at http://www.fire.pppl.gov.fpa_annual03.html.

5.5 INDUSTRIAL COGENERATION

5.5.1 Cogeneration Defined

Cogeneration is a highly efficient means of generating heat and electric power at the same time from the same energy source. Displacing fossil fuel combustion with heat that would normally be wasted in the process of power generation, it reaches efficiences that can triple, or even quadruple, conventional power generation. Although cogeneration has been in use for nearly a century, in the mid-1980s relatively low natural gas prices made it a widely attractive alternative for new power generation. In fact, gas-fired cogeneration is largely responsible for the decline in conventional power plant construction that occurred in North America during the 1980s. Cogeneration accounted for a large proportion of all new power plant capacity built in North America during much of the period in the late 1980s and early 1990s.

Cogeneration equipment can be fired by fuels other than natural gas. There are installations in operation that use wood, agricultural waste, peat moss, and a wide variety of other fuels, depending on local availability.

5.5.2 Siting Cogeneration Plants

Because it is impractical to transport heat over any distance, cogeneration equipment must be located physically close to its heat user. Cogeneration plants tend to be built smaller, and owned and operated by smaller and more localized companies than simple-cycle power plants. As a general rule, they are also built closer to populated areas, which cause them to be held to higher environmental standards.

In northern Europe, and increasingly in North America, cogeneration is at the heart of district heating and cooling systems. District heating combined with cogeneration has the potential to reduce human greenhouse gas emissions by more than any other technology except public transit.

5.5.3 Basic Concept of Cogeneration

Conventional power generation is based on burning a fuel to produce steam. It is the pressure of the steam that actually turns the turbines and generates power. This is an inherently inefficient process. Because of a basic principle of physics, no more than one-third of the energy of the original fuel can be converted to steam that generates electricity.

Cogeneration, in contrast, makes use of the excess heat, usually in the form of relatively low-temperature steam exhausted from the power generation turbines. Such steam is suitable for a wide range of heating applications, and effectively displaces the combustion of carbon-based fuels, with all their environmental implications.

In addition to cogeneration, there are a number of related technologies which make use of exhaust steam at successively lower temperatures and pressures. These are collectively known as combined-cycle systems. They are more efficient than conventional power generation, but not as efficient as cogeneration, which normally produces about 30% power and 70% heat. Combined cycle technologies can be financially attractive despite their lower efficiencies, because they can produce proportionately more power and less heat.

Cogeneration equipment recaptures the exhaust and water heat of power generation equipment and converts it into hot water, steam, space heat, process heat, air conditioning, and many other useful purposes. No additional fuel is used to provide this source of “free” energy. Figure 5-65 illustrates the process by which this “free” energy is captured and utilized.

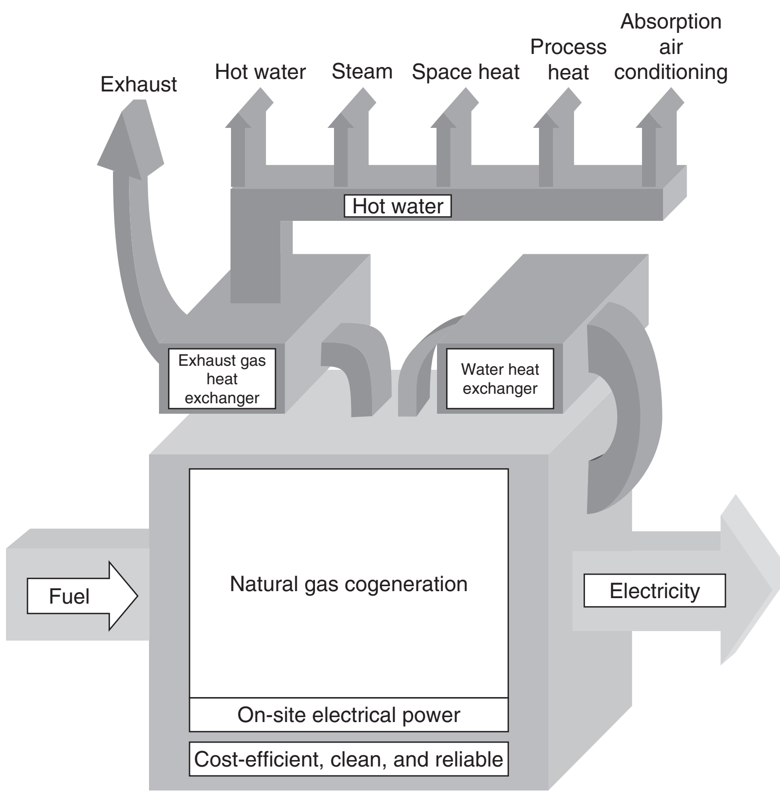


FIGURE 5-65 Diagram of the cogeneration process.

5.5.4 Advantages of Cogeneration

Cogeneration technology provides greater conversion efficiencies than traditional generation methods since it captures heat that would otherwise be wasted. This can result in more than doubling the thermal efficiency.

By recycling the waste heat, cogeneration systems achieve electrical efficiencies of 50% to 70%; a dramatic improvement over the average 33% efficiency of conventional fossil-fueled power plants. Higher efficiencies reduce air emissions and leading greenhouse gases, which are associated with climate change.

Environmental Issues. Carbon dioxide emissions will also be substantially reduced. Cogeneration systems predominantly use natural gas, a fuel source which emits less than half the greenhouse gas per unit of energy produced than the cleanest available thermal power station.

It is claimed by the Australian Cogeneration Association that cogeneration systems are more environmentally friendly, flexible, efficient, and can be more cost effective than traditional systems—especially when network costs and losses are taken into account.

Trigeneration provides even greater efficiency than cogeneration. Trigeneration is the conversion of a single fuel source into three useful energy products: electricity, steam or hot water, and chilled water. Trigeneration converts and distributes up to 90% of the energy contained in the fuel burned in a turbine or engine into usable energy. This is excellent efficiency compared to the 30% conversion that is typical for the standard electric generator.

5.5.5 Where Is Cogeneration Being Used?

A 1996 survey conducted by the International Cogeneration Alliance showed that the Netherlands was the world's leader of cogeneration use closely followed by two other Scandinavian countries—Denmark and Finland.

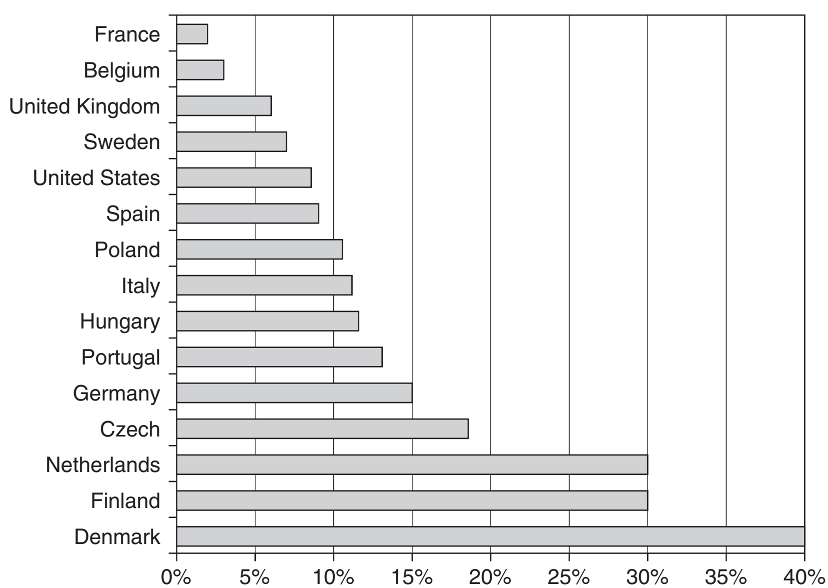


FIGURE 5-66 Percentage of national power production generated by cogeneration systems. ***Source:*** American Council for an Energy-Efficient Economy.

In 1900, half the electricity generated in the U.S. came from plants that also provided industrial steam or district heating. By the 1970s, cogeneration had fallen to less than 5%, but interest in this technology is being renewed.

The U.S. Department of Energy (DOE) set an initiative to double the use of combined heat and power systems in the United States by 2010. The U.S. National Energy Policy recommended that the president direct the U.S. Environmental Protection.

Agency (EPA) to take an active role in promoting cogeneration. While progress is slow, some states, along with EPA and DOE, are trying to lower regulatory hurdles for combined heat and power systems.

Figure 5-66 shows the percentage of national power production generated by cogeneration systems in 1997 in a variety of European countries, along with the United States.

As older power plants age and need to be replaced, and as competitive pressures to cut costs and reduce emissions of air pollutants and greenhouse gases increase, owners and operators of industrial and commercial facilities will be actively looking for ways to use energy more efficiently.

BIBLIOGRAPHY

Clayton, Sarah, 1995. *Natural Gas-Fired Cogeneration Study*, Victoria, B.C., Canada, Oil and Gas Policy Branch.

Devine, Michael D., et al., 1987. *Cogeneration and Decentralized Electricity Production*, Boulder, CO., Westview Press.

Herlock, J. H., 1987. *Cogeneration—Combined Heat and Power (CHP): Thermodynamics and Economics.* 1st ed., New York, Pergamon Press.

SECTION 6
PRIME MOVERS

Former contributors: William H. Day, Donald H. Hall, and Lawrence R. Mizin.

CONTENTS

6.1 STEAM PRIME MOVERS

6.1.1 Steam Engines and Steam Turbines

Steam prime movers are either reciprocating engines or turbines, the former being the older, dominant type until 1900. Reciprocating engines offer low speed (100 to 400 r/min), high efficiency in small sizes (less than 500 hp), and high starting torque. In the Industrial Revolution, they powered mills and steam locomotives. Steam turbines are a product of the twentieth century and have established a wide usefulness as prime movers. They completely dominate the field of power generation and are a major prime mover for variable-speed applications in ship propulsion (through gears), centrifugal pumps, compressors, and blowers. Single steam turbines can be built in greater capacities (over 1,000,000 kW) than any other prime mover. Turbines offer high speeds (1800 to 25,000 r/min) and high efficiencies (over 85% in larger units); require minimum floor space with relatively low weight; need no internal lubrication; and operate at high steam pressures (5000 lb/in^2 [gage]), high steam temperatures (1050°F), and low vacuums (0.5 inHg [abs]). Steam turbines have no reciprocating mass (with resulting vibrations) nor parts subject to friction wear (except bearings), and consequently provide very high reliability at low maintenance costs.

6.1.2 Steam-Engine Types and Application

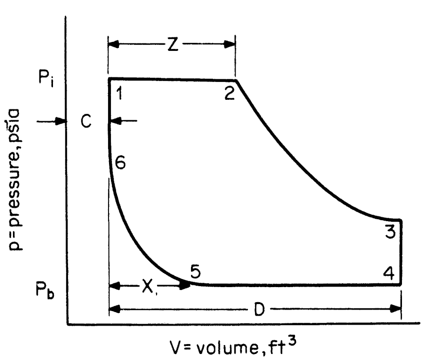
The former great diversity in engine types has been reduced so that (1) simple D-slide engines (less than 0.100 hp) are used for auxiliary drive and (2) single-cylinder counterflow and uniflow engines (less than 1000 hp), with Corliss or poppet-type valve gear, are used for generator or equipment drive in factories, office buildings, paper mills, hospitals, laundries, and process applications (where noncondensing by-product power operations prevail). Multiple-expansion, multicylinder constructions are largely obsolete except for some marine applications. Although engines as large as 7500 kW have been built and are still found in service, the field is generally limited to engines less than 500 kW in size. Engine governing is by flyball or flywheel types to (1) throttle steam supply or (2) vary cutoff.

6.1.3 Steam-Engine Performance

The basic thermodynamic cycle is shown in Fig. 6-1. The net work of the cycle is represented by the area enclosed within the diagram and is represented by the mean effective pressure (mep), that is, the net work (area) divided by the length of the diagram. The power output is computed by the "plan" equation:

$$\text{hp} = \frac{p_m Lan}{33{,}000} \tag{6-1}$$

where hp = horsepower; p_m = mep, pounds per square inch; L = length of stroke, feet; a = net piston area, square inches; and n = number of cycles completed per minute.



Where	Usual value
D = displacement	0.05–20 ft^2
C = clearance	0.03–0.2
Z = cutoff, fraction of D	0.1–0.6
X = compression, fraction of D	0.1–0.8
p_i = initial pressure	100–300 lb/in^2 (abs.)
p_b = back pressure	2–30 lb/in^2 (abs.)
p_m = mean effective pressure	50–125 lb/in^2 (abs.)

FIGURE 6-1 Pressure-volume diagram for a steam-engine cycle. Phase 1-2, constant-pressure admission at P_i; phase 2-3, expansion, $pv = C$; phase 3-4, release; phase 4-5, constant-pressure exhaust pipe at P_b; phase 5-6, compression, $pv = C$; phase 6-1, constant-volume admission.

The theoretical mep and horsepower are larger than the actual indicated values and are customarily related by a diagram factor ranging between 0.5 and 0.95. The shaft or brake mep and horsepower are lower still, with mechanical efficiency ranging between 0.8 and 0.95.

6.1.4 Steam Turbines—General

1. *Expansion of steam through nozzles and buckets*. Basically, steam turbines are a series of calibrated nozzles through which heat energy is converted into kinetic energy which, in turn, is transferred to wheels or drums and delivered at the end of a rotating shaft as usable power.

2. *Impulse, reaction, and Curtis staging*. Turbines are built in two distinct types: (1) impulse and (2) reaction. *Impulse turbines* have stationary nozzles, and the total stage pressure drop is taken across them. The kinetic energy generated is absorbed by the rotating buckets at essentially constant static pressure. Increased pressure drop can be efficiently utilized in a single stage (at constant wheel speed) by adding a row of turning vanes or "intermediates" which are followed by a second row of buckets. This is commonly called a *Curtis* or 2-row stage.

In the *reaction* design, both the stationary and rotating parts contain nozzles, and an approximately equal pressure drop is taken across each. The pressure drop across the rotating parts of reaction-design turbines requires full circumferential admission and much closer leakage control.

To illustrate the variations in energy-absorbing capacities of an impulse stage, a 2-row impulse stage, and a reaction stage, one must start with the general energy equation as applied to a nozzle:

$$\frac{V_1^2}{2gJ} + H_1 = \frac{V_2^2}{2gJ} + H_2 \tag{6-2}$$

which is reduced to

$$\text{Jet velocity, ft/s} = 223.7\sqrt{\Delta H} \tag{6-3}$$

where V_1 is assumed to be zero, and ΔH is the enthalpy drop (isentropic expansion) in Btu per pound as obtained from the Mollier chart for steam (Fig. 6-2).

Assuming a typical wheel pitch line speed (W) of 550 ft/s and initial steam conditions of 400 lb/in^2 (abs.), 700°F (H_1 = 1363.4 Btu/lb), the optimum energy-absorbing capacities of each type can be derived.

Table 6-1 illustrates that the energy-absorbing capability of the Curtis stage is 4 times that of an impulse stage and 8 times that of a reaction stage.

Because of this capability, the 2-row Curtis stage has found many applications in the process industries for small mechanical-drive use (up to 1000 hp) where the inlet steam can be taken from one process header and the exhaust steam sent out to a lower-pressure process header. As energy costs increase, however, the lower efficiency attainable with these small-volume-flow single-stage units offsets some of the desirable features (e.g., speed control, low cost, etc.). All modern turbines over 1000 hp are multistage for good efficiency, varying from 3 to 4 stages on noncondensing units with a small pressure ratio up to 20 or more stages on large reheat condensing units. Reaction (Parsons) designs generally have more stages than impulse (Rateau) designs. All large units have an impulse (1- or 2-row) first stage because there is no pressure drop on the moving rows, which makes it more suitable for partial-arc admission.

3. *The control stage.* The first stage of the turbine must be designed to pass the maximum flow through the unit at rated inlet steam conditions. The pressure required at less than rated flow will decrease if the nozzle area is held constant, resulting in a throttling loss through the control valves of the unit at partial flows. Very early in the development of steam turbines, it was recognized that if full throttle pressure could be made available to the first-stage nozzles across the load range, the maximum isentropic energy that would be available for work and overall efficiency would be increased at part load. Most first stages now use sectionalized first-stage nozzle plates with 4, 6, or 8 separate ports (depending on steam conditions, unit size, and manufacturer).

P-V-T Values for Dry Saturated Steam

Pressure, in.Hg abs	Temp., °F	Sp. vol., ft³/lb
0.5	58.8	1250
1.0	79.0	652
1.5	91.7	445
2.0	101.1	339
3.0	115.1	232
Pressure, lb/in² (abs.)		
14.7	212.0	26.8
50	281.0	8.52
100	327.8	4.43
200	381.8	2.29
300	417.3	1.543
400	444.6	1.161
600	486.2	0.770
900	532.0	0.501
1200	567.2	0.362
1800	621.0	0.218
2400	662.0	0.141
3206	705.4	0.050

NOTE: 1 in = 25.4 mm; $t_{°C} = (t_{°F} - 32)/1.8$; 1 ft³/lb = 0.0624 m³/kg.

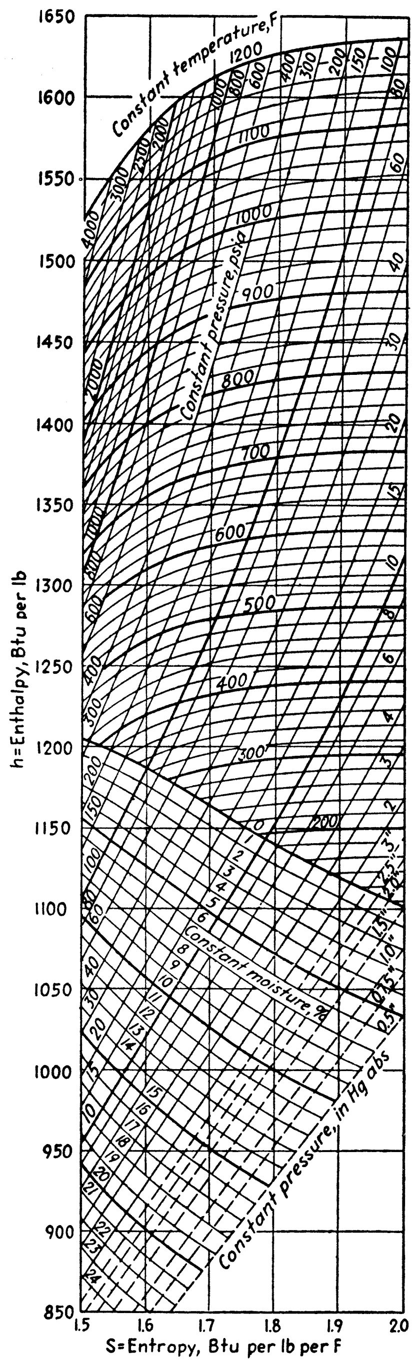


FIGURE 6-2 Mollier chart for steam (ASME steam tables.)

TABLE 6-1 Energy-Absorbing Capability of the Curtis Stage

Stage type	Impulse	2-row wheel	Reaction Sta.	Reaction Rot.	Combined
Theoretical W/V for peak efficiency	0.50	0.25	. . .	. . .	0.707
Required jet velocity, ft/s	1100	2200	550	550	
ΔH required, Btu/lb	24.2	96.6	6.045	6.045	12.09
Required P_2 lb/in^2 (abs) from Mollier chart	327	168	380.6	361.9	361.9
Stage P_1/P_2	1.224	2.38	1.052	1.105	1.105

Note: 1 ft/s = 0.3048 m/s; 1 lb/in^2 = 0.06895 bar; 1 Btu/lb = 2.326 kJ/kg.

The flow to each port is controlled by its own valve, and the valves are opened sequentially. As each valve is opened to its governing point, the full throttle pressure (minus stop-valve and control-valve pressure loss) becomes available to the arc of nozzles fed by that valve. The overall result is a greater availability of energy to do work.

4. *Steam-path design.* Condensing-turbine sizes increase with the development of longer last-stage buckets and, consequently, the last-stage dimensions (length and diameter) are the first to be determined; these dimensions fix the diameter of the L-1 stage and the optimum energy (pressure drop) which can be placed on that stage. This stage in turn defines the parameters of the L-2 stage and so on up to the first stage, and it can be said that steam paths of turbines are designed backward except for the first stage. In the 1970s, the largest-capacity single-flow condensing turbine was approximately 120,000 kW. Larger ratings are obtained by multiplying the number of exhaust stages (usually the last 5 to 7 stages are involved) by 2, 4, 6, or 8 times to satisfy the rating requirements. This practice is limited to the larger blades to round out a product line to well over 1,000,000 kW.

6.1.5 Turbine Efficiency

1. *Nozzle and bucket.* The turbine stage efficiency is defined as the actual energy delivered to the rotating blades divided by the ideal energy released to the stage in an isentropic expansion from P_1 to P_2 of the stage. The most important factors determining the stage efficiency are the relationship of the mean blade speed to the theoretical steam velocity, the aspect ratio (blade length/passage width), and the aerodynamic shape of the passages. Figure 6-3 describes the typical variation in nozzle and bucket efficiencies with velocity ratio and nozzle height.

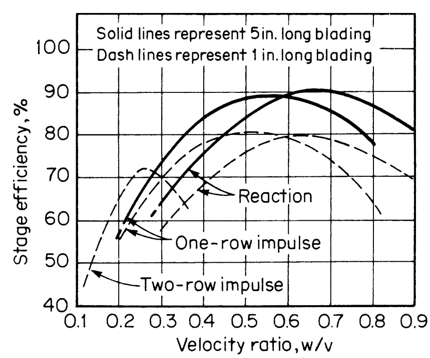


FIGURE 6-3 Approximate relative efficiencies of turbine stage types.

2. *Losses*

Clearance leakage. A 100% efficiency cannot be obtained because of friction in the blading and clearance between the stationary and rotating parts, and because the nozzle angle cannot be zero degrees. Axial clearance increases in the stages further from the thrust bearing to satisfy the need to maintain a minimum clearance at extreme operating conditions when the differential expansion between the light rotor and heavy casing is at its worst. To reduce this leakage, radial spillbands are used. These thin, metal-strip seals may be attached to the diaphragm or casing and extend close to the shroud bands covering the rotating blades. This clearance can be kept quite close (0.020 to 0.060 in), and axial changes in the rotor position do not affect the clearance since the spillbands ride over the shrouds. The need to control the clearance leakage area is especially important on reaction stages with small blade heights because of the pressure drop across the moving blades.

Nozzle leakage. Leakage around the nozzles between the bore of the blade ring or nozzle diaphragm and the drum or rotor must be kept to a minimum. This leakage is controlled through the use of a metallic labyrinth packing which consists of a single ring with multiple teeth arranged to change the direction of the steam as well as to minimize the leakage area. Labyrinth packings are also used at the shaft ends to step the pressure down at the high-pressure end and to seal the shaft at the vacuum end.

Rotation loss. Rotation of the rotor consist of losses due to the rotation of the disks, the blades, and shrouds. Partial-arc impulse stages have a greater windage loss within the idle buckets. Rotation losses vary directly with the steam density, the fifth power of the pitch diameter, and the third power of the rpm. In general, the windage loss amounts to less than 1% of stage output at normal rated output. At no-load conditions, windage loss for noncondensing turbines approximates 1.5% of the rating per 100 lb/in^2 exhaust pressure, and on condensing units approximates from 0.4% to 1.0% of the rating at 1.5 inHg (abs) exhaust pressure.

Carryover loss. A carryover loss (about 3%) occurs on certain stages when the kinetic energy of the steam leaving the rotating blades cannot be recovered by the following stage because of a difference in stage diameters or a large axial space between adjacent stages. Typically, this happens in control stages and in the last stages of noncondensing sections. The last stages of condensing turbines have the largest carryover losses (normally referred to as exhaust loss) because of the large variations in exhaust volumetric flow with exhaust pressure and the large variation of stage pressure ratio with load. Stages preceding the last operate with essentially a constant pressure ratio down to very low loads and consequently can be designed for peak efficiency at a wide range of loads.

Leaving loss. Condensing turbines are frequently “frame sized” by last-stage blade height. It is sometimes economical to size the unit with exhaust loss equal to 5% deterioration in overall turbine performance at the design point (valves wide-open throttle flow and 1.5 inHg [abs] exhaust pressure) when the normal expected exhaust pressure will be higher or the unit will be operating at part load for a large part of the time.

Nozzle end loss, partial arc. Control stages and partial-arc impulse stages are subject to end losses at the interface of the active and inactive portions of the blading as the stagnant steam within the idle bucket passages enters the active arc of nozzles and must be accelerated. There is also a greater turbulence in the steam jet at both ends of the active arc. In partial-arc impulse stages, the increase in efficiency due to larger blade heights (aspect ratio) is partially offset by increased rotation and end losses, and there is an optimum to this proportioning beyond which there is an overall loss.

Supersaturation and moisture loss. Moisture in the steam causes supersaturation and moisture losses in the stage. The acceleration of the moisture particles is less than that of the steam, causing a momentum loss as the steam strikes the particles. The moisture particles enter the moving blades (buckets) at a negative velocity relative to the blades, resulting in a braking force on the back of the blades. Supersaturation is a temporary state of supercooling as the steam is rapidly expanded from a superheated state to the wet region before any condensation has begun. The density is greater than when in equilibrium, resulting in a lower velocity as the steam leaves the nozzle. As soon as some condensation occurs at approximately 3.5% moisture, according to Yellot, a state of equilibrium is almost instantly achieved and supersaturation ceases.

3. *Turbine efficiency.* The internal used energy of the stage is obtained by multiplying the isentropic energy available to the stage by the stage efficiency. The sum of the used energies of all stages in the turbine represents the total used energy of the turbine. The internal efficiency of the turbine can be obtained by dividing the total used energy by the overall isentropic available energy from throttle pressure and temperature conditions to the exhaust pressure. (Note: The sum of the available energies of the stages is greater than the overall available energy and represents the reheat factor or gain attributable to the unused energy of preceding stages becoming available to following stages.) The use of overall available energy will automatically account for pressure-drop losses occurring in stop valves, control valves, exhaust hood, and piping between HP and LP elements. Other losses which must be accounted for to arrive at the turbine overall efficiency include valve-stem and shaft-end packing leakages and bearing and oil-pump losses. Determination of the overall efficiency of a turbine and its driven equipment must take into account the losses of gears or generators and their bearings as well.

6.1.6 Turbine Construction

Since the early 1900s, horizontal-shaft units have been universally used. Horizontal units may be single-shaft or double-shaft, with single, double, or triple steam cylinders on one shaft. These modern units may be throttle or multiple-nozzle governed, have one or more steam extraction points, and exhibit innumerable variations in construction. Figure 6-4 shows several of the more commonly used types of turbines in schematic cross sections.

Steam turbines may be classified into several broad categories, according to the basic purpose and design of the steam path: (1) straight condensing, (2) straight noncondensing, (3) uncontrolled extraction, (4) single, double, or triple controlled extraction, and (5) reheat. Various combinations of these features may be present in a typical unit, and occasionally unusual variations on the above types may be seen.

Figure 6-5 is a cross section of a modern automatic-extraction turbine, showing the details of construction. A steam turbine consists of the following basic parts: (1 to 3) steam path made up of rotating

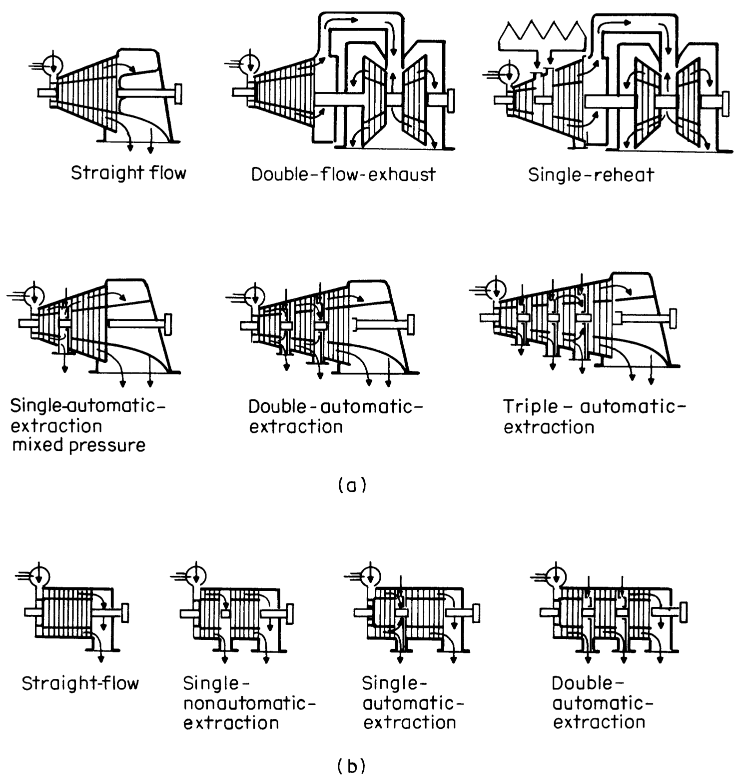


FIGURE 6-4 (a) Condensing turbines (exhaust at backpressures less than atmospheric); (b) noncondensing turbines (wide range of backpressures). (*General Electric.*)

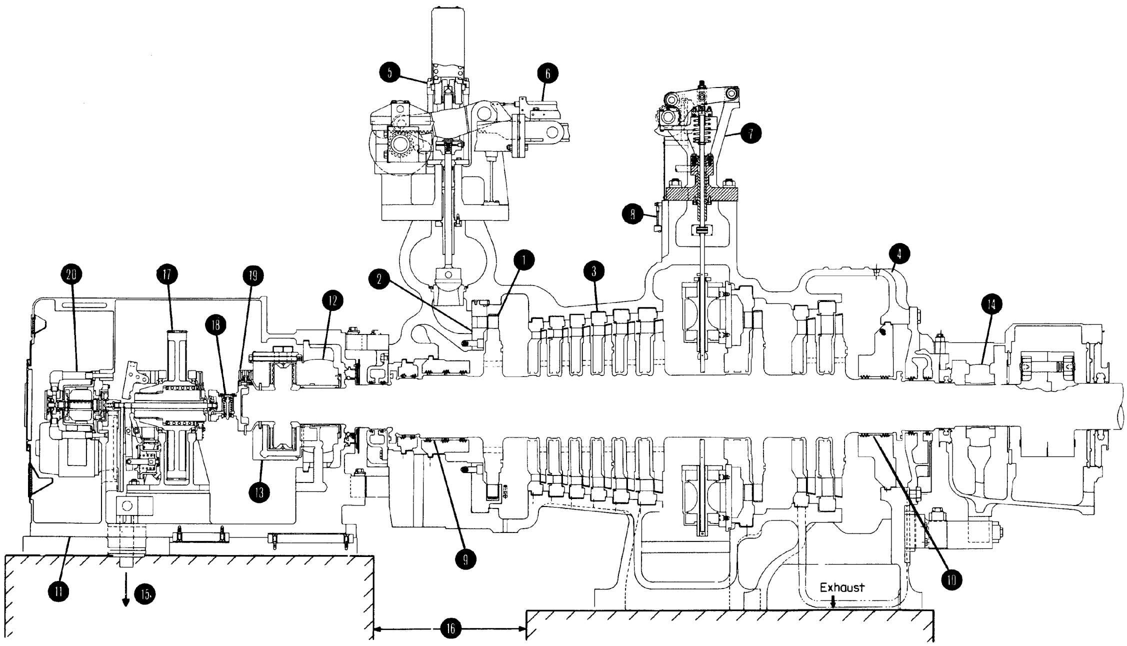


FIGURE 6-5 Cross section of a modern single-automatic-extraction noncondensing steam turbines showing construction details. (*General Electric.*)

and stationary blading (buckets and nozzles); (4) casing to contain the stationary parts and act as a steam pressure vessel; (5 to 8) controlling and protective valves, piping, and associated components to accept and control the steam admitted to the steam path; (9 and 10) packing and sealing arrangement to prevent steam from escaping into the surrounding area; (11) front standard which houses lubrication, control, and protective equipment and supports part of the casing; (12 to 14) set of journal and thrust bearings to support the rotating elements and absorb all static and dynamic rotor loads; (15) lubrication and hydraulic system for supplying bearing lubrication and (when applicable) generator seals, control, and protective oil requirements; (16) supporting foundation on which the major stationary parts rest; and (17 to 20) various accessory components, such as turning gear, control and protective components, drain valves, etc., as required by the specific application.

Turbines are constructed chiefly of carbon, alloy, and stainless steels. The rotor may be a single forging, fabricated from a shaft and separate wheels, or constructed of forged elements welded together. The buckets forming the rotating portion of the steam path are generally machined from solid stock and attached by pins, or grooves called "dovetails," to the wheels. The stationary steam path is built up of diaphragms with nozzles mounted in the heavy, two-piece casing (usually cast steel), which is bolted together on a horizontal joint. If the unit is condensing or has a low back pressure, the exhaust casing may be made up as a separate assembly and bolted to the main casing through a vertical joint. To minimize thermal stresses in high-temperature applications (950 to 1050°F), a double shell casing may be used. The rotor may sometimes be made in two pieces and coupled together, particularly in the case of the larger condensing double-flow units. A solid or flexible coupling may be used to connect the turbine rotor to its load.

Labyrinth-type packing rings, consisting of high and low teeth, are arranged at the ends of the steam path to inhibit steam from escaping into the surrounding area. (Similar packing is used at each diaphragm, and particularly at stages having control valves, to prevent excessive leakage from one stage to another within the steam path.) Associated with the external packing is a seal system which draws a vacuum to exhaust a mixture of leaking steam and air and thus prevents any steam from leaking into the surrounding room.

Bearings for supporting the turbine rotor are located in pedestals at either end and consist of journals and a thrust assembly. Normally, when steam flows through the turbine, thrust is developed in the direction of steam flow. However, unusual operating conditions or configurations often cause a thrust reversal. Therefore, it is usually necessary to provide an "active" thrust bearing for normal loading and an "inactive" thrust bearing for reverse loading.

The control-valve gear-activating equipment in a turbine usually is mounted on top of the turbine casing at the stage where steam is to be admitted. There are at least as many valve-gear assemblies as there are control stages, sometimes more if a lower valve-gear assembly is required for passing the flow. Protective or emergency valves are generally located off the machine, near the associated steam piping. Control, protective, and accessory components are often located in part in the front standard, at the pedestal housing the first journal and thrust bearing assemblies. The unit is usually supported on its foundation at the front standard and at the exhaust casing. The coupled generator shares similar foundation supports.

6.1.7 Turbine Control and Protective Systems

Steam turbines require a number of systems and components to provide control and protective capability. These may be divided into two functional categories: (1) primary control systems and (2) secondary and/or protective control systems.

Primary Control Systems. *Primary control systems* may be further subdivided into the following elements: control valves and associated operating gear, speed/load control, and pressure control. *Secondary or protective systems* consist of overspeed limiting devices, emergency valves, trip devices, and associated alarm devices.

Control-Valve Gear. Most modern turbine-generators use steam-admission control-valve designs which are as efficient as practical, in terms of pressure drop and throttling losses. Most popular are the ball-venturi valves used in the inlet stages of modern high-pressure units. The use of multiple valves, with the efficient venturi seat configuration and the tight-seating ball valves, permits partial-arc nozzle admission to the turbine with good part-load efficiency and a sequential opening action which produces nearly linear flow curves. These valves may be opened by one of two basic means: bar lift, with valves sequenced by stem lengths, or cam-operated by levers and rollers, to linearize the inherently nonlinear flow characteristics of ball-venturi valves. A common variation of the ball-venturi valve gear once widely used is the poppet-valve gear, with beveled valves and seats. This arrangement is not as efficient as a ball-venturi valve gear and is not presently very popular.

Another commonly employed valve gear, particularly on lower-pressure high-volume-flow applications, is the double-seated spool valve gear. This valve is not very efficient but passes high-volume flow and can be programmed to open in a manner similar to a ball-venturi valve.

A third scheme used somewhat in the past, but now less popular, is the grid valve, which consists of two plates with specially shaped holes arranged so that when one is rotated relative to the other, the flow area is developed as the holes coincide. High volume flow and short physical span are the grid valve's strong points, but its efficiency and accuracy are not good, and operating forces are high.

Speed/Load Control Systems. After a choice has been made from the types of valves available for steam admission to the turbine, a means must be provided for positioning these valves to obtain basic speed/load control. The primary requirement is the maintenance of an accurate, predetermined rotating speed, since all turbines are designed to operate at a specific speed or over a specific range of speeds. Every turbine, therefore, has some type of speed governor or, more generally, "speed/load control system." Its purpose is to maintain a relationship between actual turbine speed and some reference value, over a wide range of load torques.

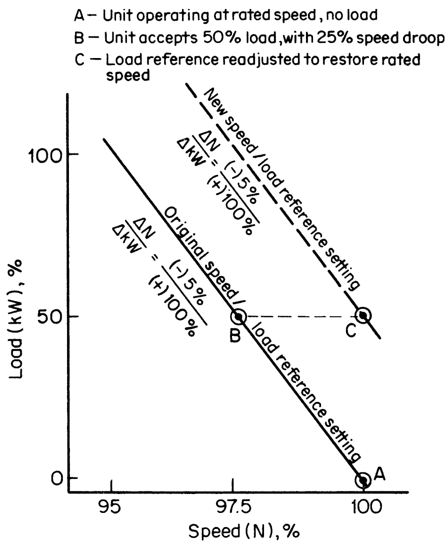


FIGURE 6-6 Steady-state speed/load regulation for a given reference setting in the speed/load system of a steam turbines. (*General Electric.*)

Land turbines used for power generation generally operate at a specific rated speed whereas marine and mechanical-drive turbines, because of the inherent coupling characteristics between rotating blades and fluids, operate over a range of speeds.

In most of the Western Hemisphere, the accepted operating frequency for turbine-generator machinery is 60 Hz. Such units having a 2-pole generator must, therefore, operate at 60 r/s, or 3600 r/min. A 4-pole unit operating at 60 Hz will rotate at half the speed of a 2-pole unit, or 1800 r/min. Most units in the United States operate at either 1800 or 3600 r/min. In most of the remainder of the world, the accepted electric frequency is 50 Hz, with common operating speeds of 1500 or 3000 r/min. In order to understand steam-turbine speed/load control, it is helpful to consider the example of constant-speed/load-based utility or industrial units.

Figure 6-6 shows the relationship designed into the speed/load control system of most such units built in the United States. The commonly accepted speed "droop," with load, for such a system is (–) 5% for 100% load change, based on a given speed/load reference setting. Note that speed droops proportionally with increasing load, on a steady-state basis. The system should be so designed that a steady-state error in speed is required to provide the command signal to move the control valve gear to accept the required load.

If such a unit operates independently, the speed/load characteristic will be as shown by the solid line in Fig. 6-6 (point *A*). If the unit is tied to a system much larger than itself, and the same system load change occurs, obviously the effect on the unit will be much less, and speed will not vary as much. The system is said to be "stiff" compared to the unit. Since speed accuracy is very important if the operating unit is isolated, any speed droop experienced with a load change (point *B*) must be corrected by changing the speed/load reference setting. This is illustrated by the dashed line of Fig. 6-6, where a 50% load change was followed by a reference correction to restore rated speed (point *C*). If an operating unit is tied to a "stiff" system, and it must accept more of the system load, a similar adjustment will cause it to pick up load with no change in speed, as the dashed line shows.

Manual speed/load reset, therefore, permits a unit, whether isolated or tied to a system, to be set to hold speed, or carry load, as the operator desires. However, if such a unit is to operate for long periods of time, and under varying load conditions, manual load reset is an inadequate solution to the problem of maintaining speed accuracy. In such cases, the use of an automatic reset device or "speed corrector" to provide isochronous control is common.

Figure 6-7 shows a very simple form of speed/load control system: a mechanical speed governor suitable for very small turbines. This type of governor uses a spring-load mechanical flyball mechanism connected to a throttling valve to directly control steam admission to the turbine. On larger units, where the forces required are too high for direct operation, a hydraulic relay governing system, as shown in Fig. 6-8, is used. In this arrangement a centrifugal flyball-type governor is connected through linkage to a double-spooled pilot valve. Oil is admitted to the pilot valve, so that when the valve moves, it ports fluid either into or out of an operating cylinder as required. The motion of the cylinder restores the pilot valve, through another linkage, to maintain a stable relationship between the pilot valve and its cylinder. Available force for operating the control valves is multiplied many times with this arrangement. On units larger than about 1000 kW, a mechanical hydraulic control system having two or more such hydraulic relays or amplifiers is used to multiply available force and to operate multiple control-valve gear systems.

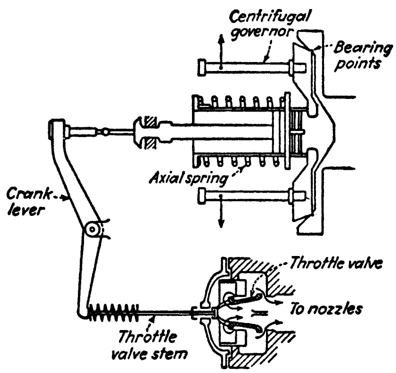


FIGURE 6-7 Mechanical governor for small turbines.

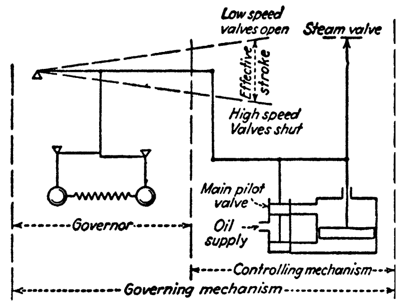


FIGURE 6-8 Governor with hydraulic power amplifier.

The electrohydraulic control system offers greater accuracy, higher operating forces, remote and centralized control capability, and more options and flexibility than any previous system. The speed/load control system in Fig. 6-9, consists of (1) a permanent magnet generator or digital-type reluctance pickup to provide a shaft-speed signal, (2) electronic circuitry for comparing the speed signal with a reference signal, (3) a high-gain servo valve to convert the resulting electric signal to a hydraulic signal, (4) a valve-gear power-actuator assembly capable of operating on high-pressure hydraulics on receipt of the servo-valve signal, (5) a feedback transducer on the power actuator to restore the servo valve to a stable condition when the desired valve position is reached, and (6) a high-pressure hydraulic system to provide the force required.

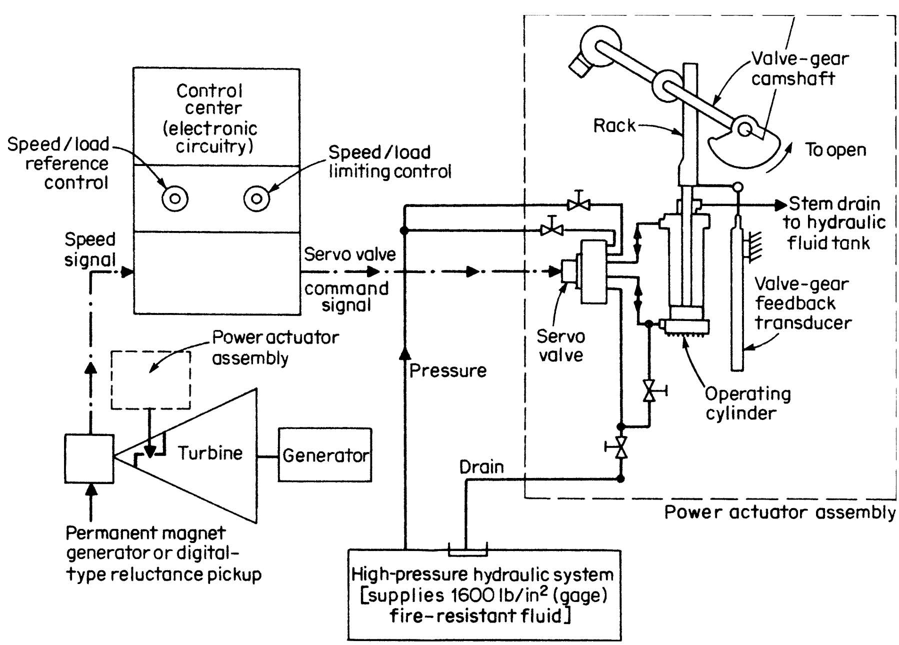


FIGURE 6-9 Schematic diagram of a basic electro-hydraulic speed/load control system. (*General Electric.*)

Other types of speed/load control systems have been applied to turbines from time to time. These include pneumatic, hydraulic, or electric devices. However, the two most common systems for turbine control are the mechanical hydraulic control (MHC) and electrohydraulic control (EHC) systems described. Another version of the EHC system was developed in the late 1960s to provide bridge control on marine turbine applications.

Pressure Control Systems. A second major area of control technology on steam turbines deals with process control. In industrial power plants particularly, it is often economical to generate and control several process flows, using steam from available steam turbines. As in the case of speed/load control, when a process is to be controlled, a definite relationship or "regulation" is established between the flow to be supplied by the turbine and the pressure. However, the possible options in process-pressure-control management are much greater than the speed/load control options described. The most common application control is for an extraction or exhaust flow from a turbine, which is to be controlled accurately in pressure and used in an industrial process. For this purpose, automatic-extraction and exhaust pressure control systems have been designed, using both the MHC and EHC technologies. Occasionally, particularly on waste-heat boiler applications, there is a need for initial pressure control as well.

Figure 6-10 is a greatly simplified schematic representation of a mechanical hydraulic control system on a single-automatic-extraction condensing turbine. The unit consists of two turbines, an HP and an LP section (each supplied by a separate valve gear), on one shaft. A flyball speed governor is used to move the two sets of valves to control speed, or load, and a bellows-type pressure governor is employed to sense process pressures and move the valves in opposite directions to control process flow and pressure. (Actual hardware required for these actions would, of course, include either mechanical hydraulic relays and linkage or electrohydraulic components.) The system is usually so designed that load and process flow variations can be satisfied at the same time with a minimum of interaction between the two variables.

Often the need exists as well for control of exhaust pressure on a noncondensing turbine-generator. In this case, since the number of variables which can be controlled is only equal to the number of control-valve stages, one variable must be sacrificed. Usually, the unit is tied to a "stiff" electrical

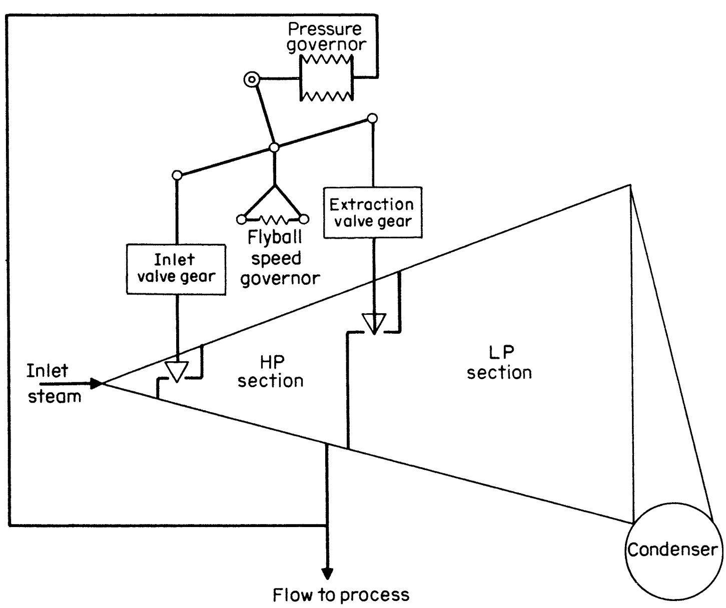


FIGURE 6-10 Simplified schematic diagram of a speed and pressure control system for a single-automatic-extraction condensing turbine. (*General Electric.*)

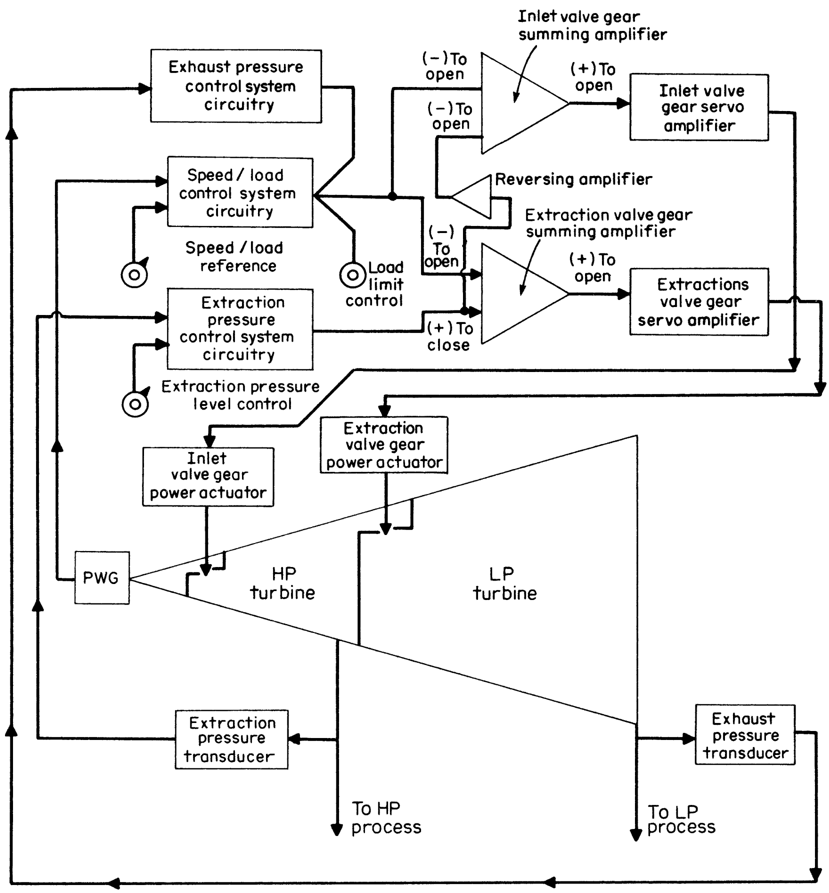


FIGURE 6-11 Schematic diagram of a modern electro-hydraulic control system for a single-automatic extraction noncondensing turbine. (*General Electric.*)

system, and speed/load control is sacrificed in favor of exhaust pressure control. Figure 6-11 shows a modern electrohydraulic control system for a single-automatic-extraction noncondensing turbine capable of controlling two process pressures for industrial needs.

6.1.8 Lubrication and Hydraulic Systems

Forced-feed lubrication of turbines and generator bearings is normally used on units above approximately 200 hp in size. On such units, the lubrication system is sometimes used to supply low-pressure seal oil for a hydrogen-cooled generator as well. Also, it often is used to supply the higher-pressure oil for the turbine control and protective systems. This is normally the case on units having a mechanical hydraulic control system and operating on turbine oil at a pressure of 250 lb/in^2 (gage) or less. On units having electrohydraulic control systems, operating at higher hydraulic pressures up to 3000 lb/in^2 (gage), and using fire-resistant fluids, a separate hydraulic-fluid power unit supplies all fluid for the control systems and usually for the protective systems as well.

6.1.9 Oil-Seal and Gas-Cooling Systems for Hydrogen-Cooled Generators

For steam turbine-generators rated up to about 40,000 kW, the electrical windings are generally cooled by air. However, above this size range, most units have hydrogen-cooled generators. Liquid cooling with hollow conductors is used on the largest units, above about 300,000 kW. Hydrogen cooling is employed because hydrogen has a thermal conductivity nearly 7 times that of air, and a density only one-fourteenth that of air. This permits reduction of windage losses and increased cooling, thereby increasing load-carrying capability for a given size of hardware.

A shaft sealing system is required to properly seal the hydrogen for cooling larger units. Oil is the sealing medium and is pressurized above hydrogen gas pressure so that it leaks across the seals to a cavity which is a receiving area for the hydrogen leaking out of the generator casing. The hydrogen-oil mixture is then scavenged through a dryer to a hydrogen control cabinet which monitors the pressure, temperature, and purity of the gas mixture, in order to maintain a safe hydrogen concentration in the generator.

6.1.10 Miscellaneous Steam-Turbine Components

In addition to the systems and components discussed in the earlier sections, steam turbines often have a number of accessory components which are important to their operation. A turning gear is provided on units rated larger than approximately 10 MW, to slowly rotate the turbine shaft before the unit is started and after it is shut down. This action helps prevent rotor bowing due to unequal heating or cooling of the rotor.

Another device of some importance is a lifting gear, for assembly and disassembly of the unit during installation and outages.

A set of turbine supervisory instruments is often included with a steam-turbine package. Typically monitored items are shaft vibration, differential thermal expansion between the casing and the rotor, expansion of the casing, eccentricity while on turning gear, thrust bearing position, speed of rotation, acceleration, control-valve position, and various other items as required by design and/or customer needs.

6.2 STEAM-TURBINE APPLICATIONS

6.2.1 Central-Station Turbines

A 60-MW, 3600-r/min *nonreheat steam turbine* is typical of those installed in smaller utility plants. Steam flows into the steam chest and through the control valves to the first-stage nozzle. After expanding through a Curtis-type, 2-row control stage, the steam flows through 16 more Rateau (impulse) stages to the exhaust. During the expansion, some steam is bled off at four or five extraction points for feedwater heating. Larger-rated units, such as those used in combined-cycle plants, require double flowing of the last five or six stages in order to provide the last-stage annulus area necessary to maintain a low leaving loss.

A 600- to 800-MW tandem-composed single-reheat steam turbine is typical of the type used in large fossil-fired central stations. Steam conditions are predominantly 2400 lb/in^2 (gage), 1000/1000°F, but some applications are at 3500 lb/in^2 (gage) and a few utilize double reheat as well. Steam enters the high-pressure turbine element through four pipes leading from the off chest-control valves to the nozzle box and the double-flow first stage. The flow is expanded through six more impulse stages before exiting from the HP casing to the reheater. The reheated steam enters at the center of the intermediate turbine and expands through seven double-flow intermediate stages before exhausting to the crossover pipe. The crossover feeds the steam to the four-flow LP elements where it is expanded to completion through six more stages before exhausting to the condenser.

A typical *nuclear steam turbine* has a capacity of 1000 to 1300 MW. Steam enters the double-flow high-pressure element at the left at 1000 lb/in^2 (gage), 546°F, and exhausts at about 200 lb/in^2 (abs) to the two large combined moisture-separator reheaters which straddle the three double-flow low-pressure elements. After the moisture is removed and the steam slightly reheated, it is passed to the six-flow

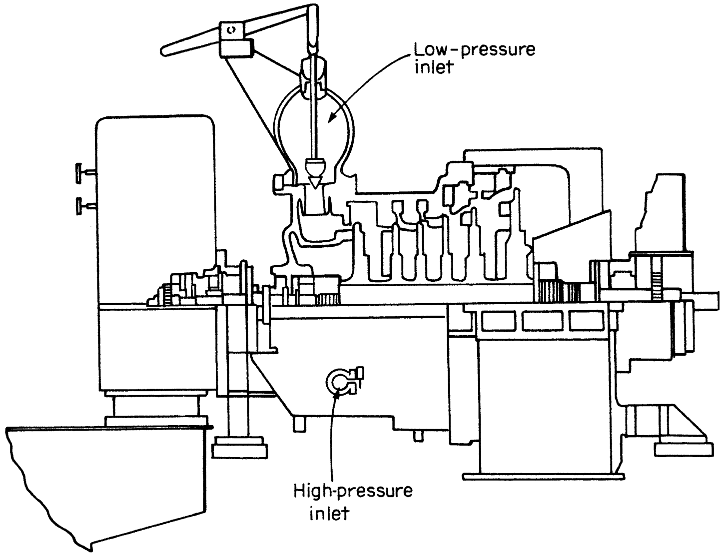


FIGURE 6-12 A 10,000 hp, 5500-r/min steam generator-steam feed-pump-drive turbine for nuclear application. (*General Electric.*)

low-pressure element where it is expanded to completion. The low superheat available with the light water reactors and the large ratings encountered require the use of 1800-r/min machinery to keep blade speeds low, reducing the erosion from moisture, and to provide the large flow areas which are more easily obtained using larger wheel diameters. The moisture-separator reheater is provided to decrease the erosion in the low-pressure elements and improve their performance. The *condensing boiler-feedwater-pump-drive turbine* was introduced in the 1960s, and became accepted rapidly because the increase in condensing annulus area served to improve both the heat rate and capacity of the station. Figure 6-12 illustrates a 10,000-hp straight condensing boiler-feed-pump-drive turbine for a nuclear application. Steam is extracted from the main unit cycle after it has gone through the moisture separator-reheater and is delivered to the inlet at approximately 150 lb/in^2 (abs) and from 0 to 100°F superheat. It is expanded to completion in the six stages. For operation at light load, steam is taken from the main steam header and sent to the high-pressure inlet in the lower half of the first stage.

6.2.2 Industrial Steam Turbines

In many industrial plants, particularly those in the pulp-and-paper, petrochemical, and related industries, the need exists for large amounts of electric power and process steam at various pressure levels. In this type of situation, industrial users can justify generating their own power and charging a large part of the cost to the process, because the steam is also needed. Plants have historically been built and expanded with various condensing and noncondensing extraction turbine types, as the process requires, and in sizes ranging from 1 to 200 MW.

6.2.3 Variable-Speed Turbines

The steam turbine is used extensively as the prime mover for ship propulsion at ratings above 10,000 shp. The cross-compound design is almost universal as it provides emergency capability for getting back to port as well as providing two pinions which divide the load on the low-speed gear, reducing gear weight. The major applications are in high-powered, high-utilization ships such as tankers and container ships. They are used almost exclusively in naval combat ships (aircraft carriers and nuclear submarines) as well as for large auxiliary supply ships.

Applications are predominantly nonreheat, but because of the steadily rising cost of fuel, reheat applications are gaining popularity.

Industry uses *mechanical-drive turbines* for a wide variety of purposes. These turbines drive paper machines; blast furnace blowers; and ethylene, ammonia, and liquefied natural-gas plant compressors because of their ability to follow the speed-output characteristics of this type of equipment without loss of efficiency from throttling, recirculation, or use of fluid couplings.

6.2.4 Special-Purpose Turbines

Turbines have been designed and built for many unusual applications. They have been built for use with working fluids other than steam in refineries and petrochemical plants. Mercury has been used as a working fluid in the *binary steam power plant.* Steam turbines have been tried in steam-locomotive applications.

Steam turbines utilizing *geothermal steam* have operated for many years in Italy, and more recently in New Zealand and the United States. The Geysers fields in California contain a high grade of geothermal energy as steam, available at the wellhead at about 100 lb/in^2 (gage). The high cost of liquid fuels has increased the attractiveness of the much more extensive geothermal brine sites as well. The geothermal-steam turbine is required to pass a much greater volume flow of steam per kilowatt generated than the central station plant. The presence of impurities in geothermal steam requires much more extensive use of alloys in the steam path and protection against moisture in the steam. The development of brine fields, which contain 300 to 500°F liquid in the wells, will require development of turbines with much larger volume-flow capacities to recover the low-level energy available. Alternative development may utilize heat exchangers which will transfer the energy to other fluids, such as isobutane, for expansion in the turbine.

6.3 STEAM-TURBINE PERFORMANCE

6.3.1 Rankine-Cycle Efficiency

The steam turbine constitutes the expansion portion of a vapor cycle, which requires separate devices, including a boiler, turbine, condenser, and feedwater pump, to complete the cycle. This vapor cycle for steam power plants is commonly called the Rankine cycle (Figs. 6-13, 6-14) and is less efficient than the Carnot cycle because the exhaust vapor is completely liquefied to facilitate pumping, and because superheat is added at increasing temperature. The work of the cycle is equal

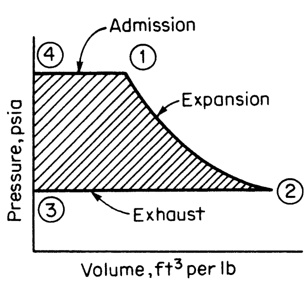


FIGURE 6-13 Pressure-volume diagram for the Rankine cycle; Phase a4-1, constant-pressure admission; phase 11-2, complete isentropic expansion; phase 2-3, constant-pressure exhaust. Crosshatched area represents the work of the cycle.

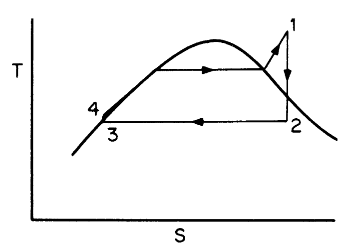


FIGURE 6-14 T-S diagram; nonreheat, nonextracting turbine cycle.

to $h_1 - h_2$ minus the small pump work $h_4 - h_3 = \nu_3(P_4 - P_3)/J$ required, and the heat added to the cycle is equal to $h_1 - h_4$. Therefore

$$\text{Rankine-cycle efficiency} = \frac{(h_1 - h_2) - \nu_3(P_4 - P_3)/J}{h_1 - h_4} \tag{6-4}$$

Cycle efficiency is not commonly used when comparing plant efficiencies because it is only indirectly determined, compared with heat rate, which can be quickly measured. In central-station practice, the station heat rate defines the heat required from fossil fuel, reactor energy, waste gases, etc., per kWh of station output (gross electric output minus all auxiliary power required within the plant). For example, a plant with a station heat rate of 10,000 Btu-fuel/kWh has a seal cycle efficiency of

$$\text{efficiency} = \frac{\text{output}}{\text{input}} = \frac{3412.1}{10{,}000} \times 100 = 34.1\% \tag{6-5}$$

Power plants for marine propulsion drive commonly use the "ships all-purpose fuel rate" (lb fuel/shp-h) as a measure of the plant's efficiency. Besides accounting for all losses required to generate the propeller-shaft output work, this fuel rate includes the requirements for the ship's "hotel" electric load, freshwater evaporators, and steam for heating and unloading cargo and cleaning tanks. A typical ship's fuel rate of 0.45 lb fuel/shp-h based on 18,500 Btu/lb fuel would indicate cycle efficiency for the ship as

$$\text{Cycle efficiency} = \frac{2544.1}{0.45 \times 18{,}500} \times 100 = 30.6\%$$

In the process industries, such as paper and petrochemical, large amounts of steam are used. Considerable by-product power can be generated by raising the boiler pressure above the process pressure and expanding the steam through a noncondensing turbine before exhausting it to the process. In this cycle, no heat is rejected because the exhaust steam is required for process and the thermodynamic cycle efficiency of this power is affected only by the boiler efficiency, the auxiliary losses chargeable to the power generation (mostly extra boiler-feed-pump work), and the mechanical and electrical losses of the turbine and generator.

The station heat rate of such by-product power generation ranges from 3900 to 4500 Btu/kWh, depending on the size of plant and the boiler efficiency, and cycle efficiencies of 80% to 85% are normal. This heat rate varies very little with turbine efficiency because energy not used to generate power is used for process. However, it is necessary to define the kilowatts generated per unit of heat to process in order to evaluate the influence of turbine efficiency or the initial steam conditions selected. Guaranteed steam rates are normally provided to evaluate the efficiency because they can be directly compared with the theoretical steam rate (TSR).

6.3.2 Engine Efficiency

The station heat rate is used to measure power plant performance, but it is of little use in evaluating the specific pieces of equipment in the cycle. The engine efficiency of the steam turbine defines its actual performance to the ideal performance. The Rankine-cycle work of the turbine is most conveniently obtained by use of the Mollier diagram (Fig. 6-15), where

$$\Delta W = h_1 - h_2 \tag{6-6}$$

and ΔW = Rankine-cycle work in Btu/lb, h_1 = steam enthalpy at throttle in Btu/lb, h_2 = steam enthalpy at exhaust in Btu/lb,

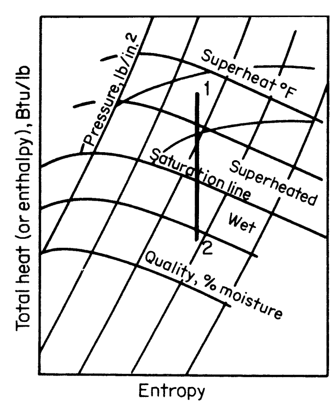


FIGURE 6-15 Steam chart (Mollier diagram.)

and h_1 and h_2 are at the same entropy (vertical line). The actual work of a real turbine is less than the ideal Rankine-cycle work, with engine efficiency defined as

$$\text{Engine efficiency} = \frac{\text{actual work, Btu/lb}}{\text{Rankine-cycle work, Btu/lb}} \tag{6-7}$$

6.3.3 Theoretical Steam Rates

The theoretical steam rate in lb/kWh for the Rankine-cycle work is expressed as

$$\text{TSR} = \frac{3412.14}{h_1 - h_2} \tag{6-8}$$

and the actual steam rate as

$$\text{ASR} = \frac{\text{TSR}}{\text{engine efficiency}} \tag{6-9}$$

The actual output of the turbine will be less than the isentropic work because of losses from nozzle and bucket friction, packing leakage, windage, carryover and exhaust losses, bearings, radiation, and throttling. The efficiency of the turbine, including its losses, is sometimes represented by the term η_{evpm}, to help identify the losses which are included in the efficiency. The subscript *e* denotes exhaust loss (significant on condensing units only), *v* denotes valve loss (control valves only), *p* denotes external packing loss (valve stems and shaft end packings), and *m* the mechanical losses. The efficiency of the generator or gear is represented by η_g.

This or similar terminology enables turbine designers to identify a particular efficiency for discussion. The stateline efficiency η represents the actual Mollier-chart expansion line of a particular turbine for given inlet steam conditions and exhaust pressure without any exhaust loss included.

The η_{evp} efficiency represents the internal efficiency corrected to include the exhaust loss, a "mean" of the control-valves pressure-drop loss, and the packing loss for the point in question. When an electric generator is the driven equipment, the overall engine efficiency η_{evpmg} is used to determine the throttle flow necessary to produce a given electric output (or vice versa).

6.3.4 Condensing-Turbine Efficiencies

Table 6-2 defines some approximate overall efficiencies of typical small straight condensing turbine-generators. The application of small turbines without regenerative feedwater heating is rare in central-station practice but is still found in waste-heat applications, process plants where feedwater heating is supplied by other sources, and increasingly in combined cycles where the stack gas is used to heat feedwater.

For turbines used in central stations, there are more satisfactory methods (see Bibliography at the end of this section) for predicting turbine efficiency, which take into account the many variables of the steam path and exhaust size and allow for inclusion of the regenerative cycle and reheat cycle in heat-balance calculations.

6.3.5 Regenerative Cycle

Steam can be extracted at several stages in the turbine to heat feedwater being returned to the boiler. In the Rankine cycle (Fig. 6-14), it was shown that the feedwater was heated from h_4 to h_1 in the boiler. By raising the temperature h_4 entering the boiler close to the saturation temperature in the drum, less fuel will be consumed in evaporating each pound of steam to h_1 conditions. The heat in the extracted steam is added to the feedwater without loss, and the heat rejected to the condenser decreases as extraction flow increases. The kilowatts do not decrease inversely with extraction flow, however, as partial expansion is made down to the extraction stages. The result is an improvement in heat rate.

TABLE 6-2 Typical Efficiencies of Straight Condensing Turbine-Generators

	Base efficiency, f_1				
kW Rating	5000	10,000	15,000	20,000	30,000
250	0.743	0.766			
400	0.733	0.757	0.769	0.777	
600	0.720	0.748	0.763	0.772	0.776
850		0.742	0.758	0.768	0.773
1250			0.754	0.765	0.770
Correction for initial superheat					
°FS	0	100	200	300	400
f_2	0.95	0.98	1.00	1.017	1.030
Correction for exhaust pressure					
P_f, inHg (abs)	1.0	1.5	2.0	3.0	
f_3	0.98	1.00	1.01	1.02	

Example: 15,000-kW turbine-generator
Steam conditions: 850 lb/in^2 (gage), 900°FTT, 2.5 inHg (abs)
TSR = 6.42 lb/kWh Superheat = 372.8°F
100% load: $\eta_{evpmg} = f_1 \times f_2 \times f_3 = 0.758 \times 1.026 \times 1.015 = 0.789$
100% ASR $= \frac{6.42}{0.789} = 8.14$ lb/kWh
Throttle flow $F_r = 15{,}000 \times 8.14 = 122{,}100$ lb/h
Source: Medium Steam Turbine Department, General Electric Co.

The actual heat rate of a regenerative cycle must be determined from a heat balance prepared by using the extraction conditions available from the turbine and the heater characteristics as specified. Table 6-3 shows the influence of heater type, temperature difference, and piping pressure drop on the gross heat rate of a 50,000-kW unit.

6.3.6 Reheat Cycle

Practically all large (over 100,000 kW) central-station plants built since 1950 have been of the reheat type. In this type of fossil-fuel plant, the steam is expanded in the turbine down to about 25% of the initial pressure, then it is sent through a reheater where it is resuperheated back up to the original initial temperature (usually 1000°F) and returned to the turbine where it is expanded to completion. In general, reheating improves the heat rate by about 5% of which only 2% is due to a higher average cycle temperature and about 3% comes from improved turbine internal efficiency due to reduced moisture and increased reheat factor.

6.3.7 Gross and Net Heat Rates

The heat-balance cycle for the turbine, condenser, feedwater pump, and heaters usually defines the gross heat rate of the cycle when a motor-driven feed pump is used. The pump work on the feedwater is included, but the power consumed by the pump is not. When a turbine-driven boiler feed pump is used in the cycle (frequently in units above 300-MW rating), the steam for the feed-pump turbine is expanded in the main unit down to the crossover to the low-pressure elements (about 100 to 200 lb/in^2 [abs]), where it is sent to the pump turbine. In these cases, the pump power required is included in the heat balance, and the heat rate calculated is called the net heat rate. In this case, only the losses from the boiler and auxiliaries (excluding the feedwater pump) must be accounted for to obtain the station heat rate.

TABLE 6-3 Heat-Rate Variation with Heater Cycle of 50,000-kW Nonreheat Turbine [1250 lb/in^2 (gage), 950°FTT, 1.5 inHg (abs)]

Base cycle HR = 8852 Btu/kWh, FWT = 430°F

	Heater no. 5	4	3	2	1	ΔHeat rate, Btu/kWh
TTD, °F	5	5	Open	5	5	Base
DCTD, °F	10	10	Htr.	10	10	Base
%Δ*P*	5	5	5	5	5	Base
Variation in terminal temperature difference, TTD						
	0	5		5	5	−10
	–5	5		5	5	−20
	5	5		5	10	+6
	10	10		10	10	+27
Variation in drain cooler temperature difference, DCTD						
	5	5		5	5	−2
	20	20		20	20	+5
	10	10		10	20	+2
Variation in pressure drop to heater, %Δ*P*						
	0	0	0	0	0	−24
	10	10	10	10	10	+26
	10	5	5	5	5	+10
Variation from drain cooled to pumped or cascaded drips						
	10	10		10	C	+12
	C	C		C	C	+30
	10	PD		10	PD	0
	10	C*		10	10	+32

*Cascaded to no. 2 heater.

Nomenclature: TTD = heater terminal temperature difference; DCTD = drain cooler temperature difference; %Δ*P* = pressure drop from turbine flange to heater shell; C = cascaded heater drains; PD = drains pumped forward.

Note: $t_{°C} = (t_{°F} - 32)/1.8$.

Source: Medium Steam Turbine Department, General Electric Co.

Table 6-4 shows typical values of gross and net heat rates at rated load for nonreheat and reheat units of typical ratings and inlet steam conditions. Exhaust pressure is 2.0 inHg (abs) in all cases, motor-driven feed-pump drive efficiency is 90%, pump efficiency is 78%, and the exhaust annulus area is normal for the rating.

The overall station heat rates of these applications are 15% to 25% greater than the net heat rate, depending on the steam-generator (boiler) efficiency and other auxiliary losses.

6.3.8 Nuclear Cycles

The turbines used in light-water nuclear cycles do not have the same freedom of steam conditions as fossil-fired cycles. The nuclear steam supply limits initial steam pressure to approximately 950 lb/in^2 (gage) at a saturated steam temperature of 540°F. The initial costs of these plants is so great that only the largest can be economically justified, and ratings are limited to about 800 MW minimum. Crossover pressures range from about 150 to 200 lb/in^2 (abs), and regenerative feedwater heating improvement is optimized by using about six heaters. These constraints and optimizations of the nuclear power-plant cycles have resulted in a rather narrow band of heat-rate fluctuations. At rated

TABLE 6-4 Typical Heat Rates of Nonreheat and Reheat Turbine-Generators

	Steam conditions			Last-stage buckets					Performance at 2.0 inHg (abs)			
Rating, MW	P_0, lb/in^2 (gage)	T_0, °F	T_{RH}, °F	No. rows	Annulus area, ft^2	Boiler feed-pump drive	Number of feedwater heaters	FFWT, °F	Gross heat rate, Btu/kWh	Net heat rate, Btu/k Wh	Throttle SR, lb/kWh	Condenser SR, lb/kWh
15	600	825		1	12	Motor	3	390	10610	10700	10.05	7.52
25	850	900		1	14	Motor	4	365	9730	9850	8.73	6.69
40	1250	950		1	19	Motor	5	444	9240	9410	8.86	6.23
50	1250	950		1	26	Motor	5	429	9080	9240	8.56	6.10
75	1250	950		1	41	Motor	5	437	8890	9040	8.44	5.96
75	1450	1000	1000	1	33	Motor	5	431	8320	8450	6.69	4.92
75	1800	1000	1000	1	33	Motor	5	448	8140	8290	6.70	4.80
100	1800	1000	1000	1	41	Motor	5	458	8110	8270	6.77	4.80
150	1800	1000	1000	2	66	Motor	6	448	7970	8130	6.51	4.66
200	2400	1000	1000	2	82	Motor	7	461	7750	7950	6.39	4.49
350	2400	1000	1000	2	132	Motor	7	473	7750	7950	6.49	4.45
350	2400	1000	1000	2	132	Turbine	7	473		7890	6.62	4.27
500	2400	1000	1000	4	222	Turbine	7	473		7830	6.57	4.26
700	2400	1000	1000	4	264	Turbine	7	473		7870	6.58	4.27
1000	3500	1000	1000	6	334	Turbine	7	504		7730	6.71	4.04

Note: 1 ft^2 = 0.0929 m^2; 1 lb/in^2 = 6.895 kPa; $t_{°C} = (t_{°F} - 32)/1.8$; 1 in = 25.4 mm.
Source: Medium Steam Turbine Department, General Electric Co.

load and 2.0 inHg exhaust pressure, the turbine net heat rate varies from about 9800 Btu/kWh at an exhaust loading of 1500 kW per square feet of annulus area up to about 10,000 Btu/kWh at 2000 kW/ft^2 of exhaust loading.

Turbines used in the high-temperature, gas-cooled reactor cycle operate at steam conditions, including reheat similar to those used in fossil-fired plants. Because of the low moisture content and lower volume flow required, 3600-r/min turbine design speed can be utilized as well. At 2400 lb/in^2 (gage), 1000/1000°F, 2.0 inHg (abs) design conditions, a heat rate of about 8300 Btu/kWh can be obtained, including the power required for gas recirculation. This represents a 15% to 17% improvement in cycle efficiency when the first plant becomes operational.

6.3.9 Combined Cycles

Improvement in the steam cycle has been rapid. The advancement of steam conditions, regenerative feed heating, reheating, and size of unit has brought the overall station heat rate down from 16,000 Btu/kWh to less than 8800 Btu/kWh on the best station.

The gas turbine developed very rapidly as a prime mover because of the improved steam cycle. During the 1950s, several exhaust-fired combined-cycle power plants were built, utilizing the exhaust gas from the gas turbine as the air supply for a fired main steam generator. After the Northeast Blackout of 1965, a large number of gas turbines were installed in the United States to serve as black start and peaking capacity units. As a result, they have become well established and accepted as a prime mover for peaking capacity.

Combined-cycle interest was renewed with the development of non-radiant-heat recovery steam generators, and the electric generation in combined-cycle plants changed from 80% to 90% steam-cycle power to 70% gas-cycle power. Since 1970, utilities have installed increasing numbers of this breed of combined-cycle plant, as they offer low initial cost, consume about one-third of the water used by straight steam plants, and provide a station heat rate 5% to 10% better than the most efficient steam plants. The major obstacle to universal acceptance of the steam-and-gas combined cycle is its fuel dependency on clean gaseous or liquid fuels.

In this combined cycle, the turbine condensate is mixed with sufficient LP economizer flow in the deaerator to raise the temperature sufficiently to avoid corrosion at the cold end of the heat-recovery steam generator. The use of regenerative feedwater heaters is avoided because of the availability of excess heat in the stack for that purpose. Part-load heat rate can be maintained at greatly reduced load in the multigas turbine plants because the units can be put in service sequentially. A so-called hockey-stick curve of part-load performance is shown in Fig. 6-16 for a combined-cycle plant utilizing four gas turbines. Below about 75% load, a gas turbine should be removed from service for best plant efficiency.

6.3.10 Noncondensing-Turbine Efficiencies

The straight noncondensing turbine generator is widely used in the process industries. The flow through a noncondensing turbine is dependent on the heat required in the process, and in order to determine the heat leaving the turbine, it is also necessary to know the enthalpy of the exhaust steam. Figure 6-17 is a plot of the mechanical and electrical efficiency of several turbine generators from 20% to 100% load. This curve permits the derivation of the internal wheel efficiency (η_{evp}) of the turbine, and ultimately the exhaust enthalpy.

6.3.11 Automatic-Extraction-Turbine Efficiencies

The automatic-extraction turbine provides the capability of delivering extraction steam at more than one process pressure simultaneously. When a condensing element is used, the kilowatt output of the unit can be maintained if the process flow varies, and will permit generation in excess of by-product power capability. The base efficiency for an automatic-extraction turbine is less than that of a straight condensing or straight noncondensing turbine because of (1) the introduction of a second control stage and accompanying parasitic losses, (2) the partial-load loss resulting when the high-pressure section of the automatic-extraction unit is passing only the nonextraction flow, and (3) the decreased pressure ratio of each section of the unit.

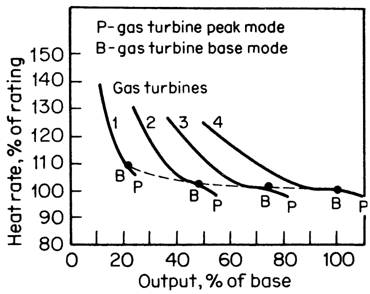


FIGURE 6-16 Output/heat rate chart for the combined-cycle power system.

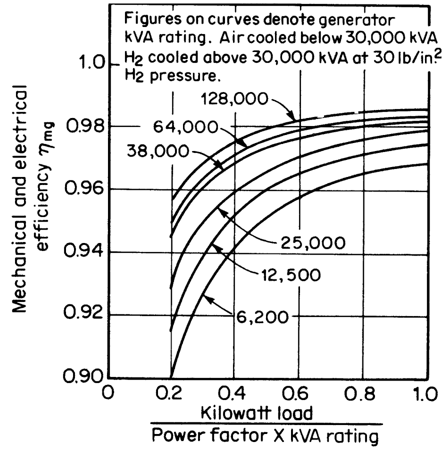


FIGURE 6-17 Typical mechanical and electrical efficiencies of turbine generators. (*General Electric.*)

The variation in throttle flow required for a change in extraction flow while maintaining constant kilowatt load can be derived from the available energy and efficiency of each section of the turbine. The extraction factor is a term which defines the relationship and represents the Δ throttle flow (ΔF_T) required for a Δ extraction flow (ΔF_x) of 1 lb/h at constant kilowatt load. The term represents a comparison of the used energy in the turbine below the automatic-extraction point to the total used energy of the turbine. The approximate extraction factor varies with the ratio of the theoretical steam rates for the turbine.

As extraction flow increases, the HP section of the unit generates more of the kilowatts and the LP section generates less until the steam flow to the LP stages is at the minimum necessary for cooling purposes. At this point, the maximum extraction for the load in question has been reached and further extraction flow must be accompanied at increasing output. The minimum cooling steam required varies with turbine size, extraction pressure, and the exhaust pressure. As an approximation, the minimum section flow in pounds per hour can be considered equal to the rating of the turbine in kilowatts.

BIBLIOGRAPHY

Cohen, H., G. F. C. Rogers, and H. I. H. Saravanamattoo. 1973. *Gas Turbine Theory*. New York: Wiley.

Division of Continuing Education. 1980. *Fundamental Principles of Gas Turbines.* Austin, TX: University of Texas.

Goldstein, Richard J. 2001. *Heat Transfer in Gas Turbine Systems*. New York: New York Academy of Sciences.

Treager, Irwin E. 1996. *Aircraft Turbine Engine Technology*. New York: McGraw-Hill.

6.4 GAS TURBINES

6.4.1 Cycles

Internal combustion engines, such as conventional automotive engines, operate on the Otto cycle; injection engines operate on the Diesel cycle; and the gas or combustion turbine operates on the Brayton cycle (Fig. 6-18), also called the *gas-turbine simple cycle*. Referring to Fig. 6-19*a*, an axial or centrifugal compressor delivers the compressed air to the combustion system, and fuel is burned

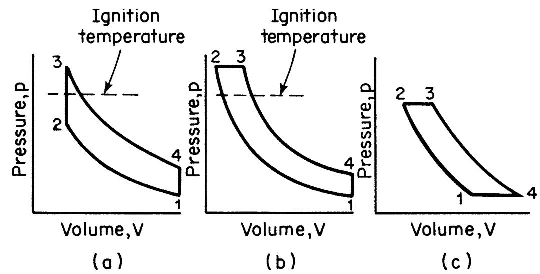


FIGURE 6-18 Ideal indicator cards (pressure-volume diagrams) for internal combustion cycles; (a) Otto cycle; (b) diesel cycle; (c) Brayton cycle. In general, phase 11-2 represents isentropic compression; phase 2-3, heat addition at constant pressure or volume; phase 3-4, isentropic expansion; and phase 4-1, heat rejection at constant pressure or volume.

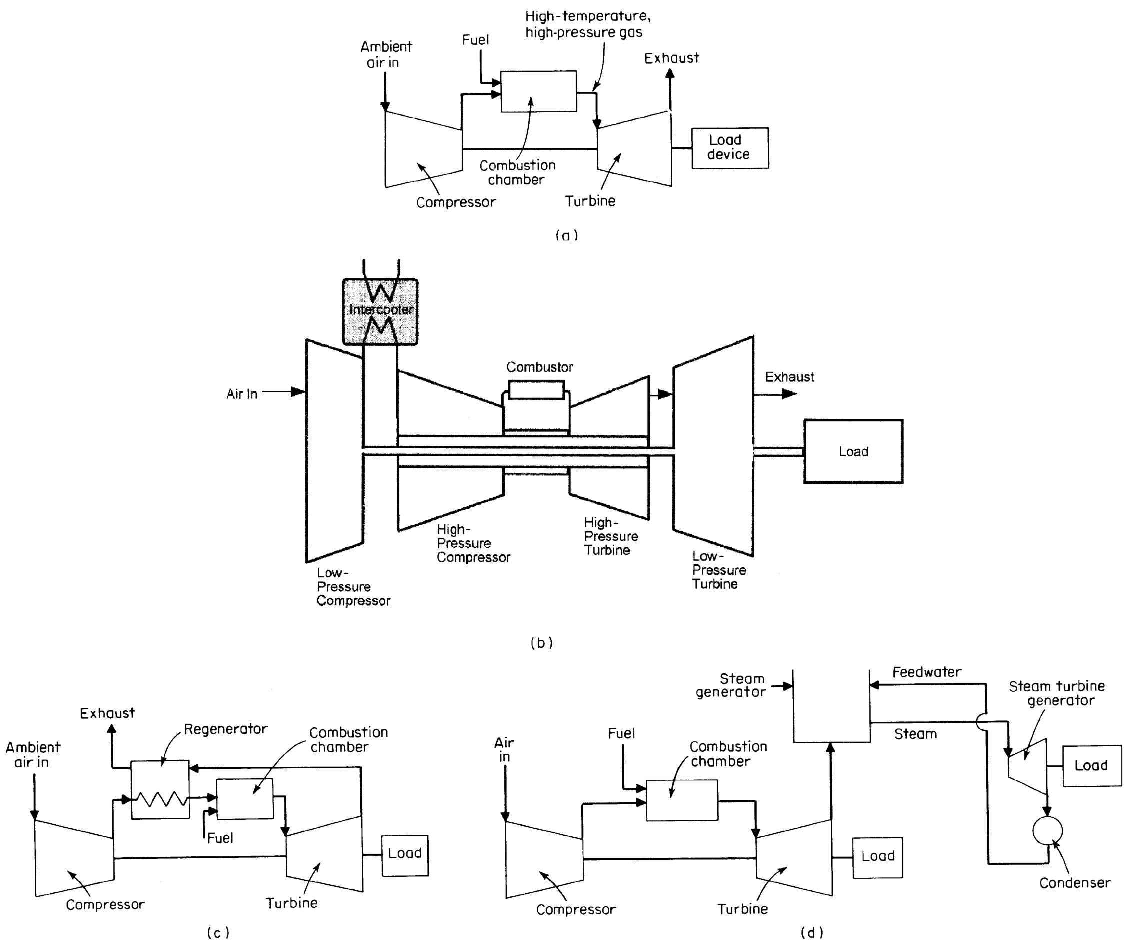


FIGURE 6-19 Typical gas-turbine cycles; (a) open, (b) intercooled (c) regenerative (d) combined.

to increase the fluid temperature. The products of combustion expand through the turbine, producing sufficient power to drive the compressor and the load. Some compressed air typically bypasses the combustor and is used to cool the turbine parts. The highest-pressure turbine airfoils contain internal cooling passages in order to maintain the metal temperature at acceptable levels for durability, while the gas path temperature is considerably higher than the metal temperature, to achieve high power and efficiency.

An improvement in power output and efficiency can be obtained through the use of an intercooler (Fig. 6-19*b*), in which air is cooled after part of the compression process. The intercooler reduces the work of compression of the high-pressure compressor and allows higher airflow and overall pressure ratio to be attained, while reducing the temperature of the cooling air for the turbine section.

Another efficiency improvement can be obtained from a regenerator or recuperator (Fig. 6-19*c*), which exchanges heat from the exhaust to the combustor inlet to reduce the fuel required to heat the gas. The highest efficiencies are available from combined cycles, where the gas turbine exhaust heat produces steam to drive a steam-turbine generator (Fig. 6-19*d*). The steam turbine output is obtained with no additional fuel input.

The most effective utilization of the fuel input is available through *cogeneration*, or combined heat and power. A typical cycle has the gas-turbine generating power and producing steam from its exhaust heat. The steam is sent to an industrial process, in some cases after generating some power in a noncondensing steam turbine. When credit is taken for the heat sent to process plus the power generated, efficiencies exceeding 80% are commonly achieved.

Simple cycles and combined cycles are by far the most commonly employed gas-turbine cycles. Some regenerative cycles were developed in the 1960s and early 1970s, but durability problems with the regenerators prevented further use. Newer regenerative cycle development started in the mid-1990s. Intercooled cycles are being studied principally as possible derivatives of commercial aircraft engines.

6.4.2 Design

Most gas turbines with outputs above 1 to 2 MW have multistage axial-flow compressors and turbines. Lower power units tend to have single-stage centrifugal compressors and radial-inflow turbines to minimize weight, size, and cost. Thermal efficiency improves with larger size, as friction and tip leakages become a smaller percentage of power produced, and multistage axial flow components become more efficient than radial stages. The same major components are included in a aeroderivative gas turbine, that is, one which was derived from an aircraft engine. This aeroderivative has two compressor sections, each driven by a separate turbine. This section of the gas turbine, called the *gas generator*, is derived from the aircraft engine. Hot gas exiting the gas generator drives the power turbine, which is connected to the load. Typical aeroderivatives differ from frame-type gas turbines in being lighter in weight and having a higher pressure ratio. Their outputs are limited to a maximum of about 50 MW because of the size of the aircraft engines from which they are derived. The largest frame-type gas turbines have outputs over 200 MW. Aeroderivatives tend to produce higher simple cycle efficiencies, whereas frame types produce the highest combined cycle efficiencies. For a given output, frame-type gas turbines tend to be somewhat less expensive than aeroderivatives.

For applications where the output shaft speed varies, as in a compressor drive, a separate power turbine is needed. This requires a multiple-shaft gas turbine (Fig. 6-20). A single-shaft gas turbine has a typical effective operating range of 85% to 105% of rated speed. With a separate power turbine to drive the load, the output shaft speed range is typically about 50% to 105% of rated speed.

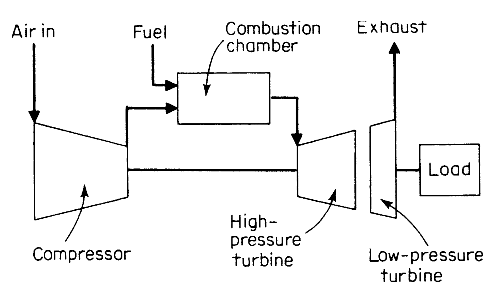


FIGURE 6-20 Multiple-shaft turbine.

6.4.3 Performance

Component efficiencies, airflow, pressure ratio, and turbine inlet (or firing) temperature are the major factors affecting gas-turbine output and efficiency. Typical multistage axial component efficiencies are in the 86% to 93% range. Material developments and turbine cooling techniques permit turbine rotor inlet temperatures to exceed 2500°F (1370°C) for the most advanced units.

6.4.4 Applications

The most familiar application of gas turbines has been for aircraft propulsion, where the turbine drives only the compressor and the remaining energy is used for thrust. The industrial gas-turbine industry started in the late 1950s, following the development of aircraft gas turbines. Generation of electric power, mechanical drive (principally gas or oil pipeline compression), and marine propulsion are the three applications of industrial gas turbines. Over 90% of the gas-turbine applications, as measured in megawatts of power produced, are in electric power generation. The majority of the navies of the world use gas turbines to propel most of their surface ships.

In electric power generation service, the combination of lower capital cost, shorter installation time, high efficiency, and environmental advantages compared to steam-turbine-based power plants has resulted in gas-turbine-based power plants having a major market for new power generation equipment.

Compliance with emissions regulations, particularly nitric oxide, or NO_x, is a major application consideration. In the late 1970s and the 1980s, water or steam was injected into the combustion reaction zone to decrease the flame temperature, which is the principal parameter affecting NO_x. Present technology is producing combustion system designs that feature premixed air and fuel in lean mixtures to reduce the flame temperature without any outside diluent.

Much work has been done to adapt gas turbines to coal fuel. Coal gasification, utilized to remove contaminants, is particularly attractive since the gas-turbine combined cycle can be integrated into the gasification process for improved efficiency. Air extracted from the gas turbine can be used as the source of oxidant for the gasification process, and the steam produced in the process can be expanded through the steam turbine. Figure 6-21 is a block diagram showing how various components relate in an integrated gasification combined cycle (IGCC).

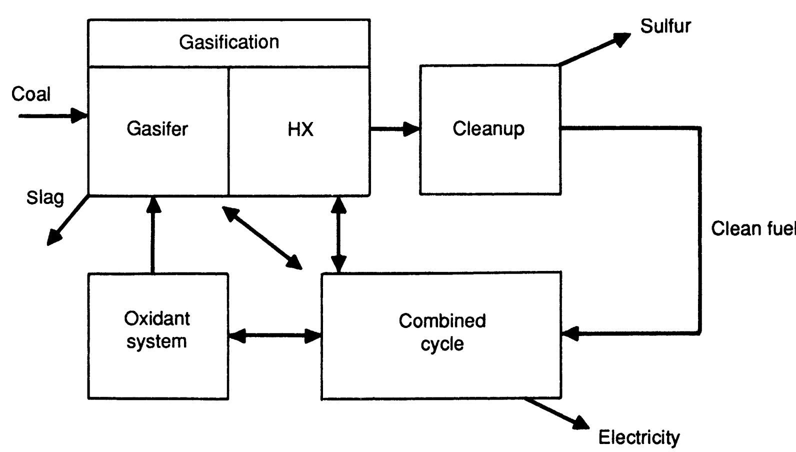


FIGURE 6-21 Possible integration among components of IGCC system.

BIBLIOGRAPHY

Bloch, Heinz P. 1996. *A Practical Guide to Steam Turbine Technology*. New York: McGraw-Hill.

Elliott, Thomas C., Kao Chen, Robert Swanskamp. 1998. *Standard Handbook of Powerplant Engineering*. New York: McGraw-Hill.

Kiameh, Philip. 2003. *Power Generation Handbook*. New York: McGraw-Hill.

2003. *Marks' Electronic Standard Handbook for Mechanical Engineers*. New York: McGraw-Hill.

Shlyakhin, P. 2005. *Steam Turbines: Theory and Design*. Honolulu, HI: University Press of the Pacific.

Woodruff, Everett B., Herbert B. Lammars, Thomas F. Lammars. 2005. *Steam Plant Operation*. New York: McGraw-Hill.

SECTION 7
ALTERNATING-CURRENT GENERATORS

John D. Amos
Samuel A. Drinkut
Aleksandar Prole
Franklin T. Emery
Lon W. Montgomery
General Engineering, Siemens Power Generation

CONTENTS

7.1 INTRODUCTION

This section deals with ac electric machines that convert mechanical power into electrical power. Such generators can be either synchronous generators or induction generators. Rotational speed of a synchronous generator is exactly at a speed that is synchronized with the ac power frequency, and this rotational speed is kept constant with varying loading conditions. Rotational speed of an induction generator is slightly above synchronous speed, and this rotational speed varies slightly with varying loading conditions. Induction generators find their major power generation application in wind turbine power generation. Synchronous ac generators dominate present-day commercial power generation by fossil fuels, nuclear reactors, and hydraulic turbines. All discussions of ac generators in this section are focused upon synchronous generators.

AC synchronous generators range in size and capability from very modest machines that are rated at a few hundred watts to the largest machines that are rated at 2000 MW. This section is intended to provide a general understanding of the nature of ac synchronous generators of this size and capability. Most discussions are focused upon larger synchronous generators with ratings above 10 MW.

This section is not intended to serve as a guide to design or manufacture of these generators, and it is not intended to serve as a textbook that explains details of the theory of function of these machines. A few textbooks about generator design and theory that may be of interest to readers of this handbook are listed in the bibliography of this section.

7.2 BASICS OF MACHINE CONSTRUCTION AND OPERATION

7.2.1 Machine Morphology

All synchronous generators function as magnetic energy conversion devices to convert mechanical power into electrical power by means of magnetic fields. The input torque provided by the prime mover (the turbine) is balanced by the magnetic torque between the stationary and rotating structures in the generator.

Several different approaches are possible to accomplish this power conversion function. For the larger synchronous generators that are primarily discussed in this section, the magnetic fields are typically established by electrical currents circulated in stationary ac windings, and rotating dc windings, and these magnetic fields are circulated within the generator through highly permeable steel structures. In such a generator, the ac winding is electrically connected to an electrical power system and physically mounted on the stationary member of the generator (the stator), and the dc winding is electrically connected to a dc power source and physically mounted on the rotating member of the generator (the rotor). Because of the prevalence of polyphase power generation, distribution, and utilization, the ac winding in all but the smallest synchronous generators is generally a polyphase winding. The most common number of phases is three.

All larger synchronous generators include an ac armature winding and a dc field winding. The electromagnetic interaction of these two windings provides the basis for ac power generation. In some of the smallest synchronous generators, with ratings below a few hundred kilowatts, the magnetic function of the dc field winding is provided by permanent magnets. In all large synchronous generators, the dc field is provided by a dc field winding. This section is limited to discussions of generators, with an ac armature winding and a dc field winding.

In most large synchronous generators, the ac armature winding is located on the stator of the machine, and the dc field winding is located on the rotor, as illustrated schematically in Fig. 7-1. An important exception is a special synchronous generator that is generally known as a brushless exciter. A brushless exciter is a relatively small synchronous generator (50 to 500 kW) that is used to provide dc electric current to the rotating field winding of a large synchronous generator. In brushless exciter, the dc field winding is mounted on the stator and the armature winding is mounted on the rotor. That said, all further discussions of morphology in this section are based upon the most common arrangement for generators of 10 MW and above, where the ac armature winding is located on the stator of the machine and the dc field winding is located on the rotor, as illustrated in Fig. 7-1.

In a generator, like that illustrated in Fig. 7-1, the magnetic circuit consists of a steel stator core that is mounted upon the steel stator case and a steel rotor that is supported on bearings that are either set into the case or separately mounted to the foundation. The coils of the armature winding are mounted in the stator core, and the coils of the field winding are mounted on the rotor. Armature winding electrical coils for generators of the type shown in Fig. 7-1 are typically deployed in radial slots formed in the inner diameter of the stator, and field winding electrical coils are typically deployed in radial slots formed in the outer diameter of the rotor, as illustrated in Figs. 7-2 and 7-3 respectively.

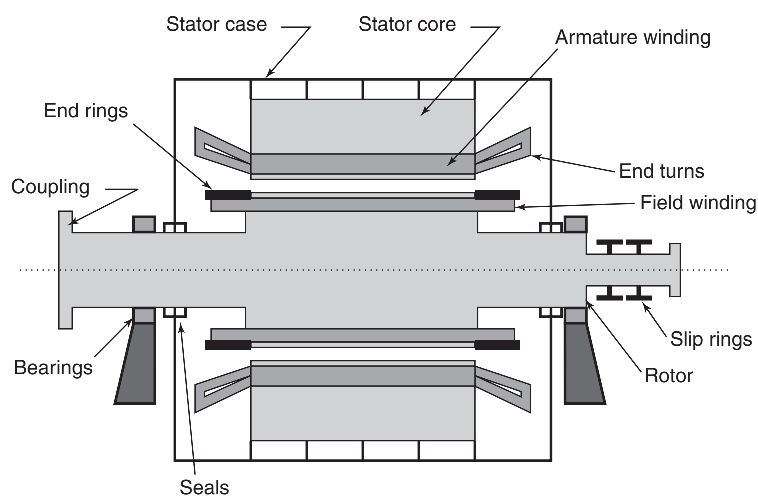


FIGURE 7-1 Elements of an ac generator.

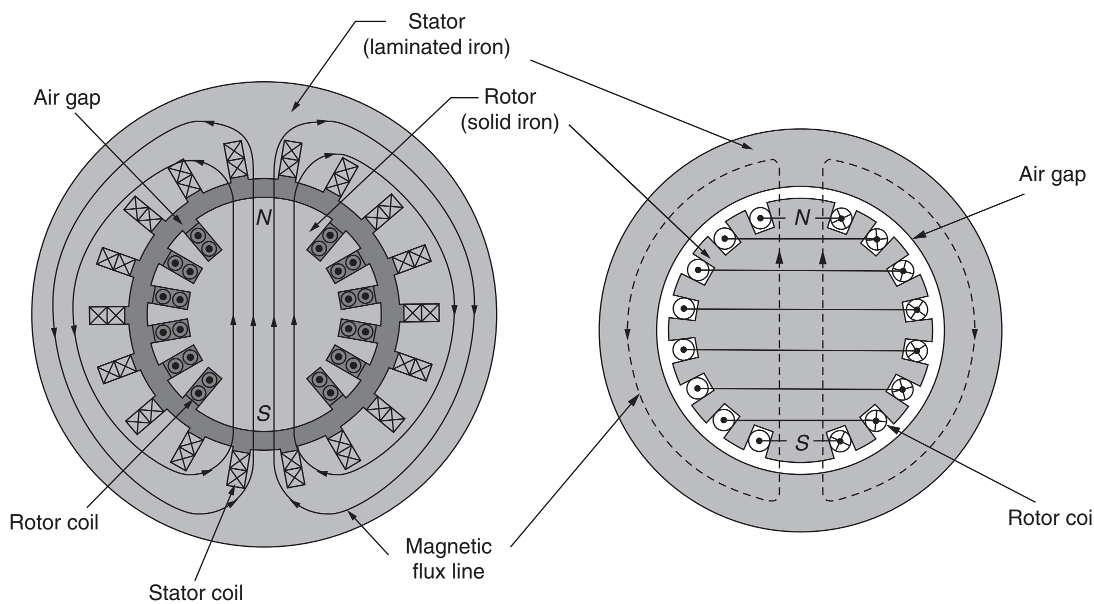


FIGURE 7-2 Round-rotor generator with two poles.

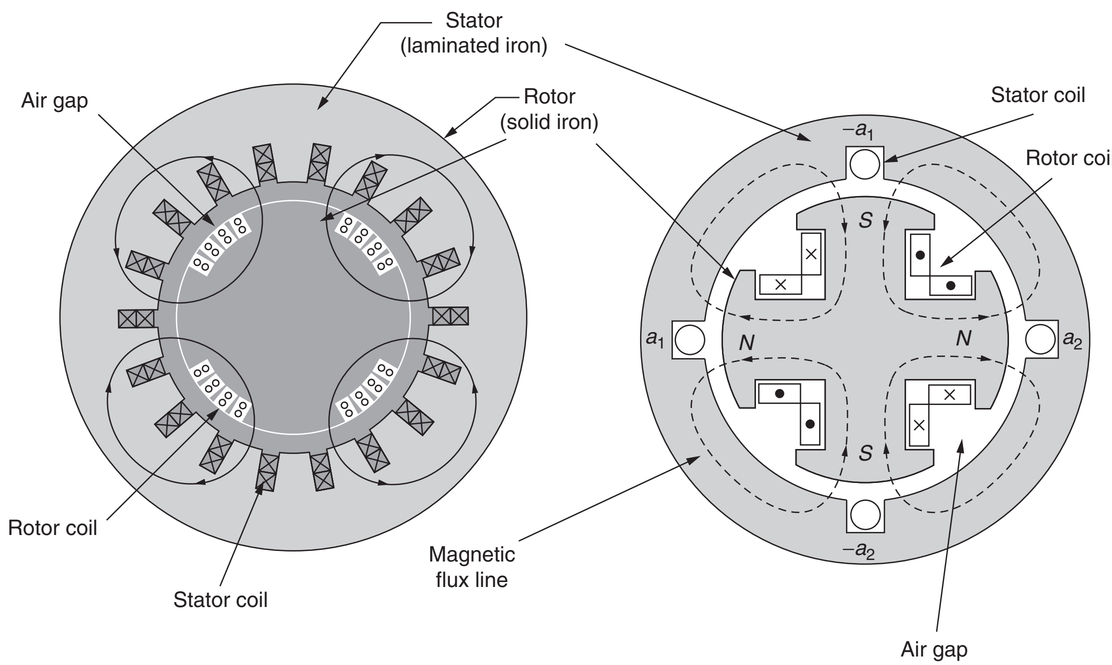


FIGURE 7-3 4-pole generator (left is round rotor, right is salient pole.

7.2.2 Poles and Frequency

The rotor and stator (field and armature) of a synchronous machine must have the same number of poles, as the magnetic interaction is between a succession of north-south magnetic-field pole pairs. The number of pole pairs for a machine will be noted as p. The relationship between electrical frequency f_e and mechanical speed N is

$$f_e = p\,\frac{N}{60} \tag{7-1}$$

where f_e is measured in hertz (Hz) and N is in revolutions per minute (r/min). A common expression for Eq. 7-1 is

$$N \cdot P = 120 \cdot f_e \tag{7-1a}$$

where P is number of poles (not number of pole pairs).

Synchronous generators are built in two elementary forms:

- *Round-rotor* machines are constructed with a rotor consisting of a cylinder of magnetic steel. In modern generators, the cylinder is formed from a single forging of vacuum degassed steel. The field winding is contained in radial slots in the surface of the rotor. Round-rotor machines usually have two or four poles as illustrated in Figs. 7-2 and 7-3 respectively. The diameter of the rotor of a typical 25-MW generator is about 700 mm. The diameter of a 2000-MW generator can approach 2 m.
- *Salient-pole* machines are constructed with a number of pole pieces mounted to a central rotor shaft. The rotor pole pieces can be solid steel or assemblies of steel plates that are bound together axially with bolts. The diameter of the rotor can range from less than 1 m in smaller salient pole generators to nearly 20 m in the largest hydroelectric generators.

In both round-rotor and salient-pole generators, the magnetic flux passing through the rotors does not vary in time, and the magnetic flux passing through the stator core does vary periodically in time at the electrical line frequency. Consequently, the rotors can be made of solid steel, but the stator cores must be made of thousands of thin layers of highly permeable electrical steel. Each layer of stator core steel is coated with a thin layer of electrical insulation.

For electric utility operation, in which generation takes place at 50 or 60 Hz, mechanical speed is inversely proportional to the number of poles. Thus, 2-pole machines, which turn at 3000 or 3600 r/min, are used for most fossil-(fuel)-fired steam turbine generators which require high shaft speeds. Most nuclear steam turbine generators, which have a lower shaft speed requirement, employ 4-pole designs and therefore turn at 1500 or 1800 r/min. Turbine generators for both fossil and nuclear power plants are typically round-rotor designs.

Hydroelectric generators, which typically have much lower shaft speeds than turbine generators and consequently require a large number of poles, are generally built as salient-pole machines. This is true also for generators intended for operation with large reciprocating engines, such as medium-speed diesels.

7.2.3 Basis of Operation

A synchronous generator works by causing an interaction of two multiple-pole magnetic-field distributions, those of the stator (armature) and, rotor (field). The interaction is said to be *synchronous* because, if the rotor is turning at the speed described by Eq. (7-1), the armature and, rotor magnetic fields are turning at the same physical speed. The synchronous operation may be described in two elementary ways, referred to as the magnetomotive force (mmf) *method* and the *flux method*. These are described here, assuming a simple, linear, round-rotor model for the machine. It should be noted that this model will, of necessity, be modified later to fully understand operation of the machine.

MMF Method. A principal feature of a synchronous generator is the mutual inductance between phases. Assuming a 3-phase machine, the mutual inductances between the field winding and the 3-phase windings are

$$M_{af} = M \cos p\phi \tag{7-2}$$

$$M_{bf} = M \cos\left(p\phi - \frac{2\pi}{3}\right) \tag{7-3}$$

$$M_{cf} = M \cos\left(p\phi + \frac{2\pi}{3}\right) \tag{7-4}$$

where M is the peak value of mutual inductance and ϕ is the angle between the axes of the field winding and the stator phase winding designated a. If it is further assumed that phase-phase inductances, both self- and mutual, are not a function of rotor position, the use of energy methods gives a simple expression for machine torque:

$$T = -pMi_a i_f \sin p\phi - pMi_b i_f \sin\left(p\phi - \frac{2\pi}{3}\right) - pMi_c i_f \sin\left(p\phi + \frac{2\pi}{3}\right) \tag{7-5}$$

If the rotor turns at a constant angular velocity $\omega/p = 2\pi f_e/p$, the field current is held constant at a value of I_f and the three stator currents are sinusoids in time, with the same amplitude and with phases that differ by 120°

$$p\phi = \omega t + \delta_i \tag{7-6}$$

$$i_a = I \cos \omega t \tag{7-7}$$

$$i_b = I \cos\left(\omega t - \frac{2\pi}{3}\right) \tag{7-8}$$

$$i_c = I \cos\left(\omega t + \frac{2\pi}{3}\right) \tag{7-9}$$

torque is

$$T = -\frac{3}{2} pMII_f \sin \delta_i \tag{7-10}$$

Note that torque is proportional to the product of the two current amplitudes and to the sine of the phase angle between the current distributions. Further, the torque is acting in a direction so as to align the two current distributions.

Flux Method. The flux method for estimating machine torque focuses on voltage (and hence flux) induced in the machine stator. If L_a is phase self-inductance and L_{ab} is phase-phase mutual inductance, flux linked by armature phase a is

$$\lambda_a = L_a i_a + L_{ab} I_b + L_{ab} I_c + MI_f \cos p\theta \tag{7-11}$$

Noting that the sum of phase currents is, under balanced conditions, zero and that the mutual phase-phase inductances are equal, this is

$$\lambda_a = (L_a - L_{ab}) i_a + MI_f \cos p\theta = L_d i_a + MI_f \cos p\theta \tag{7-12}$$

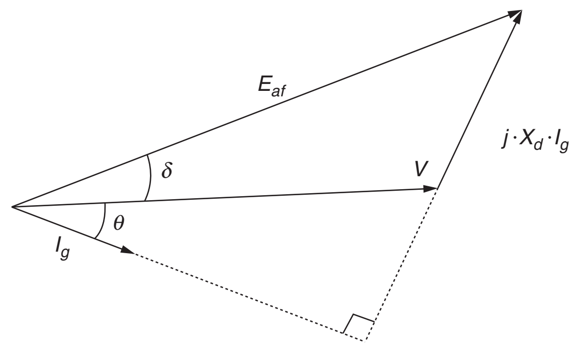
where L_d denotes synchronous inductance.

This flux is described by the equivalent circuit of Fig. 7-4, where

$$E_{af} = j\omega MI_f e^{j\delta} \tag{7-13}$$

and δ is the phase angle between *internal* voltage E_{af} and *terminal* voltage V, and $X_d = \omega L_d$.

Assume $R_a << X_d$, where R_a is the armature resistance.



(*a*) Overexcited (lagging power factor).

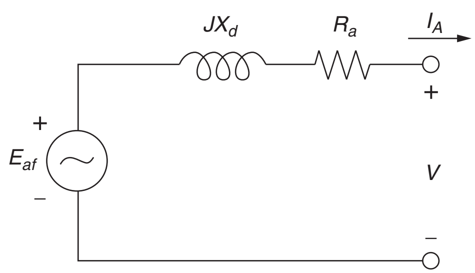
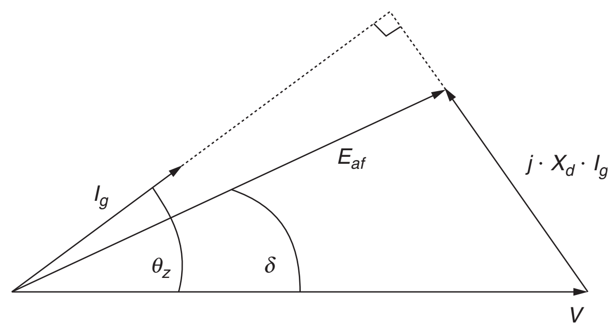


FIGURE 7-4 Steady-state equivalent circuit (R_a is neglected for analysis).



(*b*) Underexcited (leading power factor).

FIGURE 7-5 Round-rotor synchronous generator.

If the generator is connected to a voltage source (i.e., if V is fixed), terminal current is

$$I = \frac{V - E_{af}e^{j\delta}}{jX_d} \tag{7-14}$$

Real and reactive power into the terminals of phase a are

$$P_a = -\frac{1}{2}\frac{VE_{af}}{X_d}\sin\delta \tag{7-15}$$

$$Q_a = \frac{1}{2}\frac{V^2}{X_d} - \frac{1}{2}\frac{VE_{af}}{X_d}\cos\delta \tag{7-16}$$

Considering all three phases, total generated power is

$$P = -3P_a = \frac{3}{2}\frac{VE_{af}}{X_d}\sin\delta \tag{7-17}$$

Phasor diagrams illustrating the operation of a round-rotor synchronous generator are shown in Fig. 7-5. When the machine is *overexcited*, terminal current *lags* terminal voltage. When the generator is *underexcited*, terminal current *leads* terminal voltage.

7.2.4 Salient-Pole Machines: Two-Reaction Theory

Salient-pole generators, such as hydroelectric generators, have armature inductances that are a function of rotor position, making analysis one step more complicated. The key to analysis of such

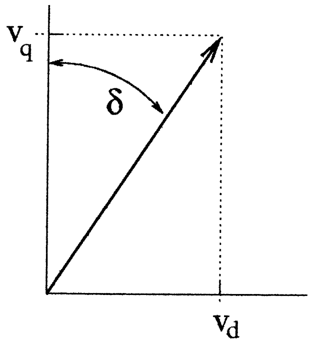


FIGURE 7-6 Direct- and quadrature-axis voltages.

machines is to separate mmf and flux into two orthogonal components. The two components are aligned with the direct axis and the quadrature axis of the machine (Fig. 7-6). The direct axis is aligned with the field winding, while the quadrature axis leads the direct by 90°. Then, if ϕ is the angle between the direct axis and the axis of phase *a*, flux linking phase *a* is

$$\lambda_a = \lambda_d \cos\phi - \lambda_q \sin\phi \tag{7-18}$$

Then, in steady-state operation, if $V_a = d\lambda_a/dt$ and $\phi = \omega t + \delta$, we obtain

$$V_a = -\omega\lambda_d \sin\phi - \omega\lambda_q \cos\phi \tag{7-19}$$

or

$$V_d = -\omega\lambda_q = V\sin\delta \tag{7-20}$$

$$V_q = \omega\lambda_d = V\cos\delta \tag{7-21}$$

One might think of the *voltage* vector as leading the *flux* vector by 90°. If the machine is linear, fluxes are given by

$$\lambda_d = L_d I_d + MI_f \tag{7-22}$$

$$\lambda_q = L_q I_q \tag{7-23}$$

Note that, in general, $L_d \neq L_q$, and for wound-field machines, $L_d > L_q$. Terminal voltage now has these components

$$V_d = -\omega\lambda_q = -\omega L_q I_q = V\sin\delta \tag{7-24}$$

$$V_q = \omega\lambda_d = \omega L_d I_d + \omega MI_f = V\cos\delta \tag{7-25}$$

which is easily inverted to produce

$$I_d = \frac{V\cos\delta - E_{af}}{X_d} \tag{7-26}$$

$$I_q = -\frac{V\sin\delta}{X_q} \tag{7-27}$$

where $X_d = \omega L_d$, $X_q = \omega L_q$, and $E_{af} = \omega MI_f$.

In the complex frame of reference

$$V = V_d + jV_q \tag{7-28}$$

$$I = I_d + jI_q \tag{7-29}$$

complex power is, in the sense of a generator

$$P + jQ = -\frac{3}{2}VI^* = -\frac{3}{2}\{(V_d I_d + V_q I_q) + j(V_q I_d - V_d I_q)\} \tag{7-30}$$

or

$$P = \frac{3}{2}\left[\frac{VE_{af}}{X_d}\sin\delta + \frac{V^2}{2}\left(\frac{1}{X_q} - \frac{1}{X_d}\right)\sin 2\delta\right] \tag{7-31}$$

$$Q = -\frac{3}{2}\left[\frac{V^2}{2}\left(\frac{1}{X_q} + \frac{1}{X_d}\right) - \frac{V^2}{2}\left(\frac{1}{X_q} - \frac{1}{X_d}\right)\cos 2\delta - \frac{VE_{af}}{X_d}\cos\delta\right] \tag{7-32}$$

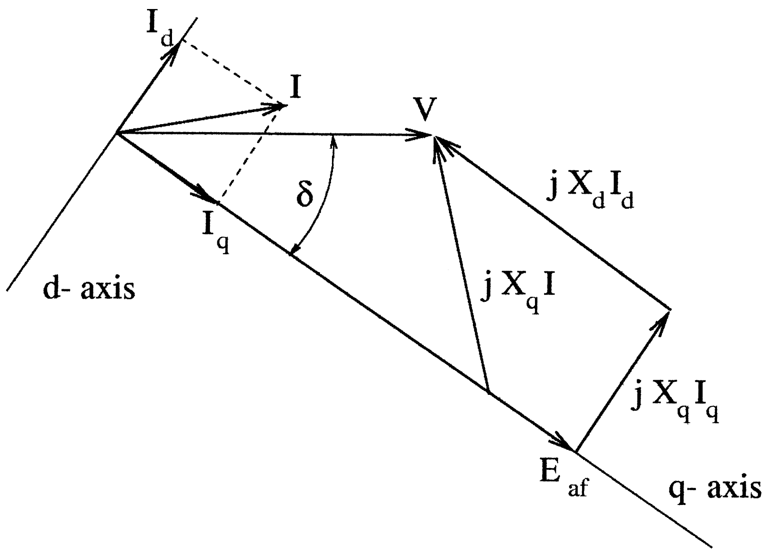


FIGURE 7-7 Vector diagram for salient-pole machine.

Figure 7-7 shows a phasor diagram for a machine with "positive" saliency (and ignoring stator resistance). It is helpful to note that in such a machine, a vector with complex amplitude $j\underline{I}X_q$ begins along the quadrature axis and ends at the ends of the terminal voltage vector.

7.2.5 Machine Size and Utilization

Generators produce torque through interaction between magnetic flux density and current over the surface of the stator, and reaction torque through the same type of interaction over the surface of the rotor. The stator and rotor face each other across the air gap. Power produced is

$$P = \omega_{\text{mech}} T = 2\pi \frac{f}{p} T = 2\pi \frac{N}{60} T \tag{7-33}$$

where torque produced is

$$T = 2\pi R^2 l \sigma \tag{7-34}$$

where f = electrical frequency, Hz
N = mechanical speed, r/min
R = stator inner radius
l = active length
σ = average value of air gap shear stress, given approximately by

$$\sigma \approx \frac{1}{\sqrt{2}} B_1 K_z \tag{7-35}$$

where B_1 is the peak value of fundamental magnetic flux density at the stator surface and K_z is the effective surface current density root mean square (rms) of the armature. The effective surface current density is ampere-turns per unit of periphery, modified by pitch and distribution factors, and by power factor.

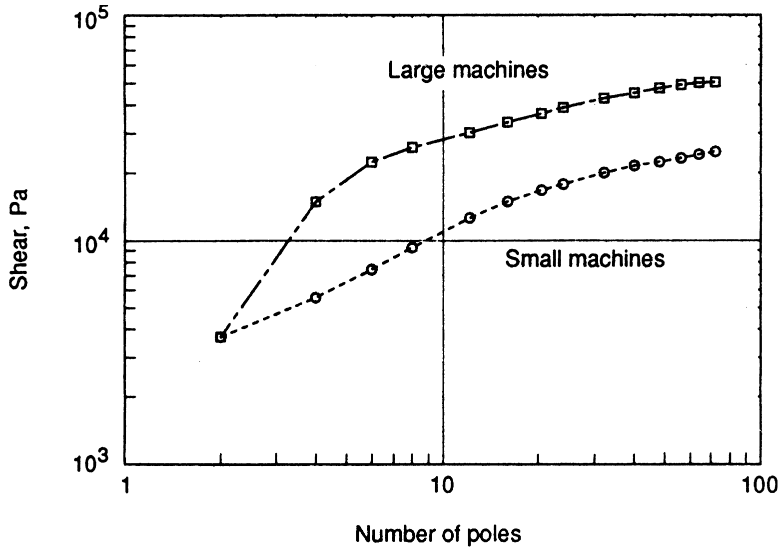


FIGURE 7-8 Typical shear stress, salient-pole, air-cooled generators.

Shear stress normally increases with pole pitch for a particular voltage and number of poles because the deeper armature slots and greater field coil space allow more ampere-conductors per unit of periphery. Typical shear stress levels for indirectly cooled, salient-pole generators are shown in Fig. 7-8. Shear is higher for directly cooled machines and a consequence of increased current density, as shown in Fig. 7-9.

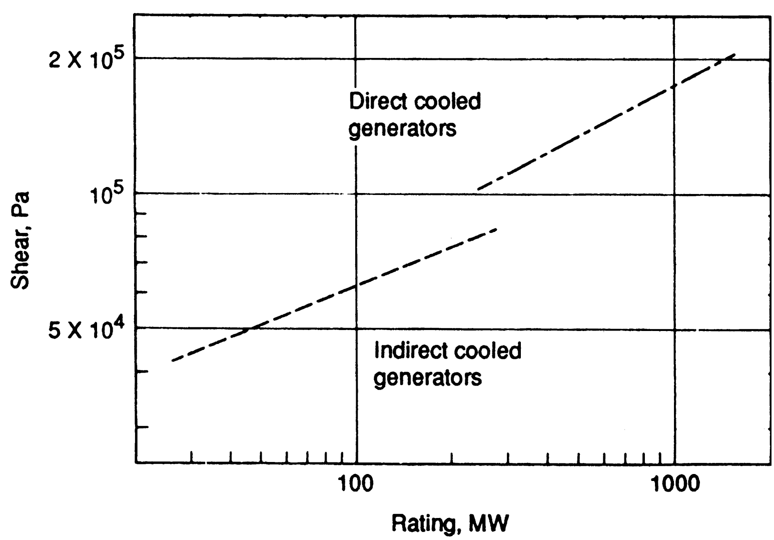


FIGURE 7-9 Typical shear stress, high-speed generators.

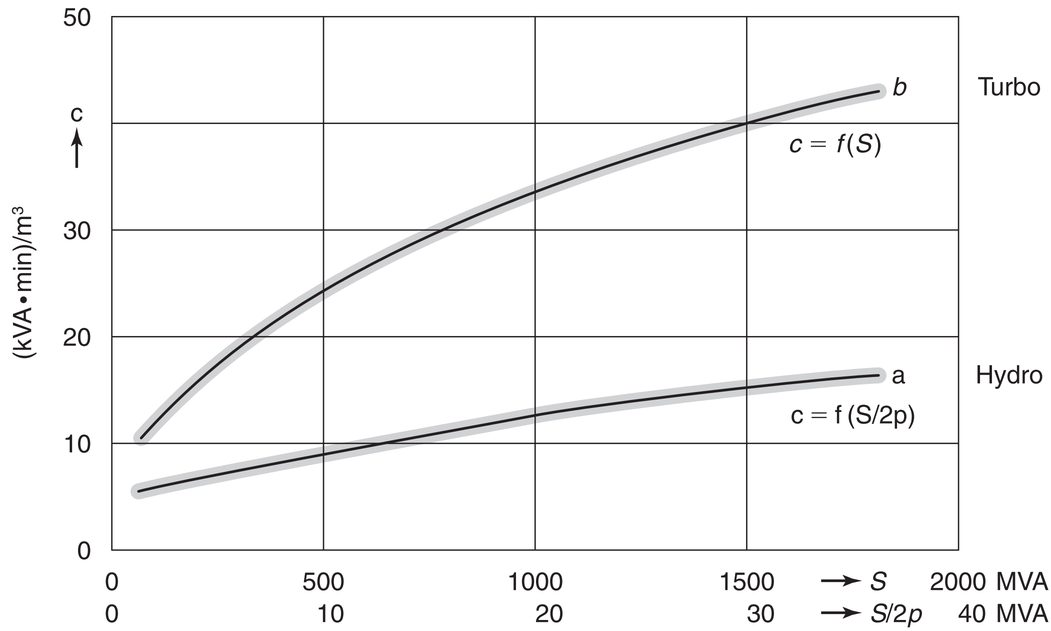


FIGURE 7-10 Utilization factors. (*ABB.*)

Another rough estimate of machine size employs the *utilization factor:*

$$c = \frac{S}{D^2 l N} \tag{7-36}$$

where S is machine rating, D and l are outside diameter and length, respectively, and N is rotational speed. Typical values for the utilization factor c are shown in Fig. 7-10. Here the utilization factor depends on unit rating (high-speed machines) or unit rating per pole (low-speed machines) and on cooling.

7.3 ELECTROMAGNETICS

Sketches of the major elements of synchronous generators, together with field "flux forms," are shown in Figs. 7-11 and 7-12. Here the generator is "unrolled" and shown as if the air gap were flat. Note that the main flux distribution air gap is not exactly sinusoidal. Thus the voltage produced will contain time harmonics. Assign C_n to be the amplitude of the nth-space harmonic of the flux density, relative to the peak amplitude B_0.

The fundamental flux per pole is

$$\Phi_1 = \frac{2RlB_0C_1}{p} \tag{7-37}$$

7.3.1 Generated Voltage

The rms amplitude of generated voltage is, for the time, fundamental

$$V_1 = \frac{1}{\sqrt{2}}\,\omega N_a k_w \Phi_1 \tag{7-38}$$

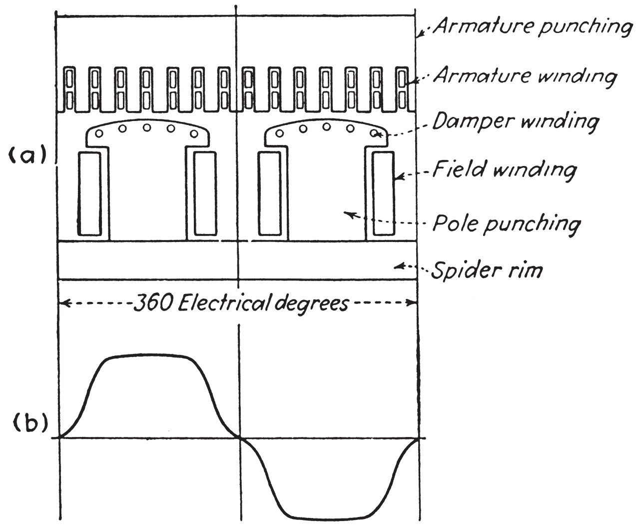


FIGURE 7-11 (*a*) Generalized sketch of one pair of poles for a salient-pole machine; (*b*) flux form for typical pair of poles (current in field winding only).

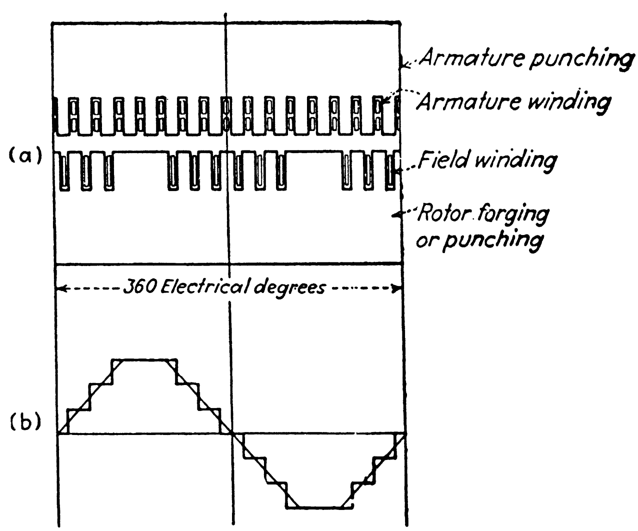


FIGURE 7-12 (*a*) Generalized sketch of one pair of poles for cylindrical rotor machine; (*b*) flux form for a typical pair of poles (current in the field winding only).

where V_1 is the time fundamental voltage induced in a phase and N_a is the number of turns in the phase. In a 3-phase machine connected in wye, line-line voltage is $\sqrt{3}$ times phase voltage. The winding factor k_{wn} is explained in Eqs. (7-40) through (7-42).

Harmonics. Time-harmonic voltages will be induced, and they are

$$V_n = V_1 \frac{C_n}{C_1} \frac{k_{wn}}{k_w} \tag{7-39}$$

It should be noted that certain time harmonics, those referred to as the *triplen* harmonics (odd factors of 3), are induced in all three phases of a 3-phase machine in phase. Thus, if line-line voltage is measured, these harmonics (orders 3, 9, 15, . . .) will turn out to have zero amplitude.

Tooth Ripple Effects. The voltage waveform predicted in Eq. (7-39) neglects the effect of armature teeth and slots on the air-gap flux. The resulting modulations of the air-gap fields dc not generate voltages in armature conductors directly, rather induce currents in the field winding, any damper windings that may exist, and in the rotor body itself and slot wedges. These currents produce flux components, which in turn produce voltages in the armature winding. The orders of these voltages are the integer multiple of the number of slots per pole pair, ±1. Thus, in a winding with 24 slots per pole pair, the harmonics will have orders of 23, 25, 47, 49, 71, 73, and so on.

Estimation of voltage waveform is actually quite complex. While Eq. (7-39) may be used for a first-order estimate, manufacturers of generators typically use numerical (finite element) methods for prediction of harmonic voltage production in actual practice.

Load Effects. Load on the machine affects the harmonic content in the following ways:

1. Increased field current tends to increase all internal voltages including harmonics.

2. Armature reaction generally reduces the fundamental voltage more than the harmonics and, through tooth and slot permeance variations, introduces additional harmonics.

3. Magnetic saturation changes harmonic magnitudes.

4. The fraction of harmonic *internal voltage* that appears at the machine terminals depends on the relationship between internal and load impedance.

Voltage Waveform Standards. There are two ways of specifying the nonideal (departure from a sine wave) nature of a voltage waveform. Both of these are defined in ANSI C42.10. Limiting values for these factors are specified in ANSI C50.12, C50.13, and C50.14.

Deviation Factor. This is, as the name implies, the maximum deviation from a sine wave. It is defined as the maximum difference between the actual waveform and the equivalent sine wave, normalized to the equivalent sine wave amplitude.

Telephone Influence Factor. This is a weighted sum of the magnitudes of all harmonics in the voltage waveform. The weighting of each harmonic is intended to reflect the relative objectional effect of inductive coupling at each harmonic frequency on telephone communications.

Pitch and Breadth Factors. The winding factor is $k_{wn} = k_{pn}k_{bn}k_{sn}$. Its component parts are called the *pitch factor*

$$k_{pn} = \cos \frac{n\alpha}{2} \tag{7-40}$$

the *breadth factor*

$$k_{bn} = \frac{\sin \dfrac{mn\gamma}{2}}{m \sin \dfrac{n\gamma}{2}} \tag{7-41}$$

and, when applicable, the *skew factor*

$$k_{bn} = \frac{\sin \frac{n\theta_s}{2}}{\frac{n\theta_s}{2}}$$

where α = *pitch angle*: the electrical angle between coil halves of the armature winding (this is generally a bit less than π to reduce the impact of higher harmonics and to make armature end windings shorter)

m = number of slots per pole per phase

γ = electrical angle between slots

θ_s = electrical angle of skew

Here, the relevant angles are stated in *electrical* terms. The electrical angle is p times the physical angle. Thus, in a 4-pole machine with 72 slots, the electrical angle between slots is $2 \times 360/72 = 10°$.

In many ac generators, the skew angle is zero for which the skew factor is unity. In some cases, the stator is skewed with respect to the rotor (or vice versa) to reduce the effects of slot openings in inducing currents in rotor parts with a consequent effect on rotor surface losses and noise.

7.3.2 Example of 4-Pole, Armature-Wound Machine

As an example, consider a 4-pole armature wound in 48 slots with a 5/6 relative pitch. The *coil pitch,* or the number of slot pitches between coil halves, is then 10 slots (5/6 × 48/4). The number of slots/pole/phase is 4 (48/[2 × 2 × 3]). Then the winding factors for the first few harmonics are

Harmonic	k_p	k_d	k_w
1	0.9659	0.9577	0.9250
5	0.2588	0.2053	0.0531
7	0.2588	–0.1576	–0.0408
11	0.9659	–0.1261	–0.1218
13	–0.9659	0.1261	–0.1218
17	–0.2588	0.1576	–0.0408
19	–0.2588	–0.2053	0.0531
23	–0.9659	–0.9577	0.9250
25	0.9659	–0.9577	–0.9250
29	0.2588	–0.2053	–0.0531
31	0.2588	0.1576	0.0408
35	0.9659	0.1261	0.1218
37	–0.9659	–0.1261	0.1218

Observe that the winding factors k_w for the harmonics 23 and 25 have the same value as the fundamental voltage. This is so for all harmonics whose numbers are equal to (integer multiple of number of slots per pair of poles) ±1. These harmonics are called slot harmonics, and the armature winding cannot attenuate the voltages induced by flux waves of these orders.

7.3.3 Armature Reaction

Current in the armature conductors produces an mmf which has the same number of poles as the field structure. The fundamental harmonic of this mmf rotates at synchronous speed, and adds to the field mmf in a vector sense as shown in Fig. 7-13. For generators supplying reactive current to an inductive load, the net effect of the armature reaction is to oppose the field mmf, requiring additional field current to sustain flux.

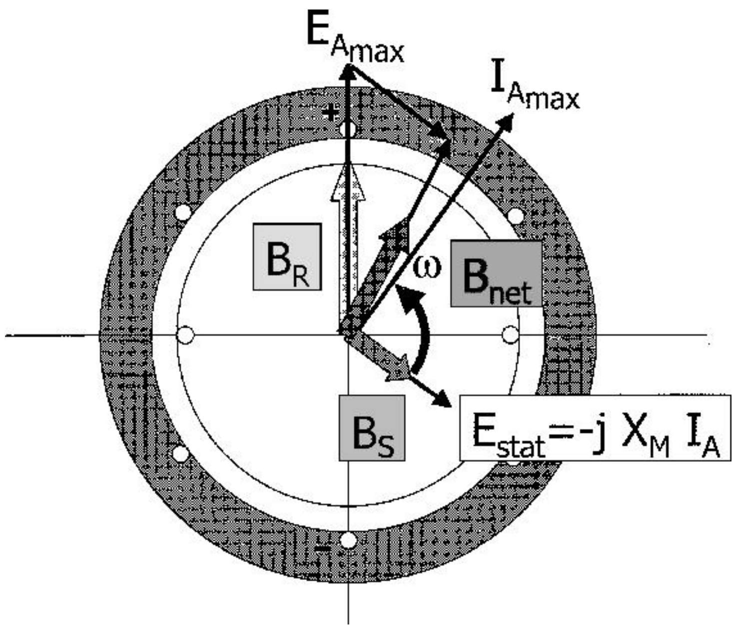


FIGURE 7-13 Armature reaction (vector analysis of magnetic fields).

The peak value of the space fundamental of armature reaction current for a 3-phase machine is

$$F_a = \frac{3}{2}\frac{4}{\pi}\sqrt{2}N_a I k_p k_b k_s \tag{7-43}$$

where I is rms terminal current, N_a is the number of series turns per phase, and the winding factors are as defined above.

Space harmonic components of armature reaction current can produce losses in the rotor as they turn asynchronously. The magnitude of the harmonic mmf of order n is

$$F_{an} = \frac{3}{2}\frac{4}{n\pi}\sqrt{2}N_a I k_{pn} k_{bn} k_{sn} \tag{7-44}$$

The rotational speed of each of these harmonics is

$$\omega_n = \mu\frac{1}{n}\cdot\omega \tag{7-45}$$

where, for a 3-phase machine, the rotational direction is same as that of the fundamental (positive) for harmonics of orders 7, 13, 19,….., and is opposite from the rotation of the fundamental (negative) for harmonics of orders 5, 11, 17, … . The electrical frequency of the pairs of these waves in the rotor frame turns out to coincide, so that the armature harmonics of orders 5 and 7 and fundamental frequency both appear in the rotor frame at 6 times the fundamental frequency.

7.3.4 Magnetic Circuit and Material

The magnetic circuit of an ac generator, as with other electric machines, is made up of the air gap, the stator teeth and backiron, the rotor poles, and the shaft section. Each of these elements has an effect on machine rating and operation.

The function of the magnetic circuit is to carry flux that links the armature conductors to produce voltage.

Air gap. The air gap constitutes the division between the rotating part of the machine—the rotor, which carries the field winding—and the stationary part of the machine—the stator, which carries the armature winding. In ac generators, the air-gap dimension is determined by the electrical characteristics of the machine. There is a trade-off between excitation mmf (toward a small air-gap dimension) and armature reaction flux (toward a large air-gap dimension). This trade-off generally results in an air gap, which is substantially larger than mechanical considerations such as machining tolerances or windage loss would dictate.

Stator Teeth and Backiron. The armature magnetic circuit carries alternating flux and is always laminated, either with complete ring laminations (for small machines) or with overlapping segmented laminations. The material most commonly used is sheet steel, of an alloy containing about 3.5% silicon, in sheets of thickness between about 0.35 and 0.65 mm. Grain-oriented steel, with reduced losses and improved permeability in the direction of rolling, is often used in large turbogenerators. Orientation in the circumferential direction is advantageous in such machines because of the large proportion of steel and moderate flux densities in the backiron. At high flux densities characteristic of the armature teeth, the advantage of grain orientation becomes less pronounced.

The active region of the armature constitutes the alternation of stator teeth and slots carrying the armature winding. The division between teeth and slots is a compromise between flux-carrying capability and current-carrying capability. The trade-off generally results in a division that is about half slots and half teeth. Flux densities in the stator teeth are usually high enough to result in moderate saturation of the magnetic material.

Rotor Iron. The magnetic flux in the rotor is nearly constant, varying in the main only slightly with changes in load and terminal voltage and with small higher frequency components due to time and space harmonics of armature flux. This allows the rotor magnetic circuit to be made of solid steel.

In turbine generators, the rotor is typically made of a single-piece forging of steel with slots for the field winding cut by machining. The losses caused by harmonic driven eddy currents in the solid steel pole faces can be problematic, and are reduced by making the air gap larger, by increasing the number of stator slots and, by choosing a suitable (short-pitch) coil throw for the armature.

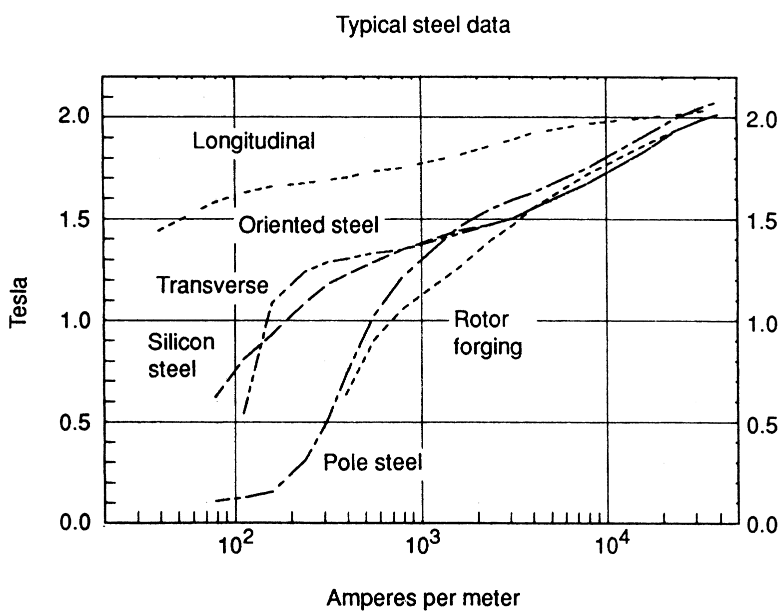


FIGURE 7-14 Magnetization curves of commonly used steels.

Salient-pole machines may have solid or laminated poles. In many cases, laminated poles are necessary to control eddy current losses. Pole laminations are commonly made of low carbon steel, 1.5 to 2 mm thick. Thinner steel, sometimes with silicon content, may be used where further control of eddy current losses is required. The shaft, or inner portion of the rotor of salient-pole machines, is often a solid forging, or in large machines such as hydroelectric generators may be fabricated from structural steel pieces.

Magnetic Materials. Typical magnetization characteristics of steel materials used in the magnetic circuit of ac generators are shown in Fig. 7-14.

7.4 MACHINE OPERATION

In normal operation, real power is dictated by the prime mover and reactive power is determined by the real power and by field current.

7.4.1 Capability Diagram

Equations (7-31) and (7-32) are approximate ways of estimating the real and reactive power output of a generator as a function of field current and torque angle. (Eq. 7-46 is provided for additional clarification.) If these are cross-plotted as shown in Fig. 7.14, a way of representing the capability of an ac generator emerges as shown in Fig. 7-15. This is called a *capability chart* for obvious reasons and four limits are shown.

$$
\begin{aligned}
P &= 3 \cdot V_\phi \cdot I_A \cdot \cos(\theta_z) = 3 \cdot V_\phi \left(\frac{E_A \cdot sin(\delta)}{X_s} \right) \\
Q &= \left(\frac{3 \cdot V_\phi}{X_s} \right) \cdot X_S \cdot I_A \sin(\theta_z) = 3 \cdot V_\phi \cdot I_A \sin(\theta_Z)
\end{aligned}
\tag{7-46}
$$

The final capability curve is shown in Fig. 7-16. Four limits are shown on this chart:

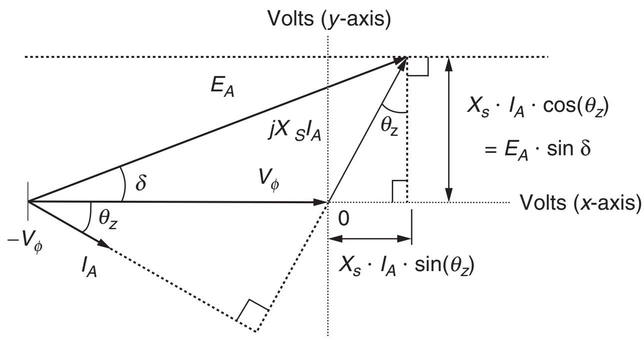
1. The field winding limit is generally related to the thermal capability of the field winding, and limits operation of the generator at high reactive power conditions.
2. The armature winding limit is generally related to the thermal capability of the armature winding, and is typically a limit on total kVA (kilovoltampere) output of the machine.
3. The stability limit is related to torque angles that are nearly at the peak of the torque-angle curve (90° for round-rotor machines).
4. Often, the configuration of magnetic flux is such that, at high reactive power absorption (negative Q), there is axial flux in the core ends, leading to excessive heating and limiting the reactive power that can be absorbed by the machine.

Very often the prime-mover power rating is plotted on the capability chart. It is, of course, a line of constant real power. The heating related limits (armature, field, and core end) may be functions of the state of the cooling system of the machine, such as hydrogen pressure. (See Sec. 7.10 on cooling.)

7.4.2 Saturation Curves and Excitation

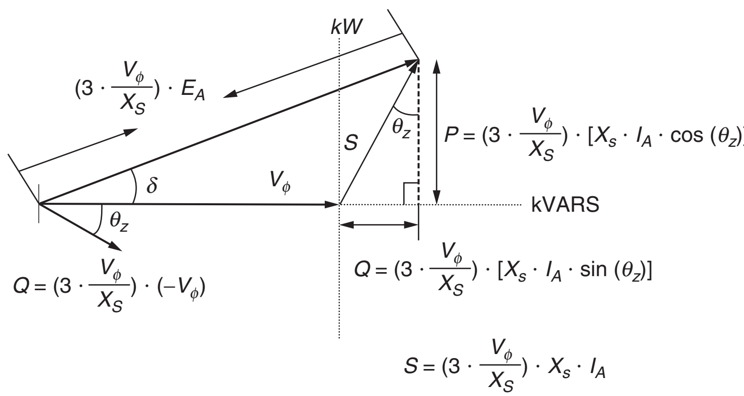
Alternating-current generators are usually operated with at least part of the magnetic circuit partially saturated, so that the linear model of machine operation implied by some of the foregoing discussion does not give exactly the right answers. What follows is an approximate way of estimating excitation requirements for an ac generator.

It should be noted that numerical methods, employing finite elements, are now available and capable of making even more accurate estimates of machine performance including excitation

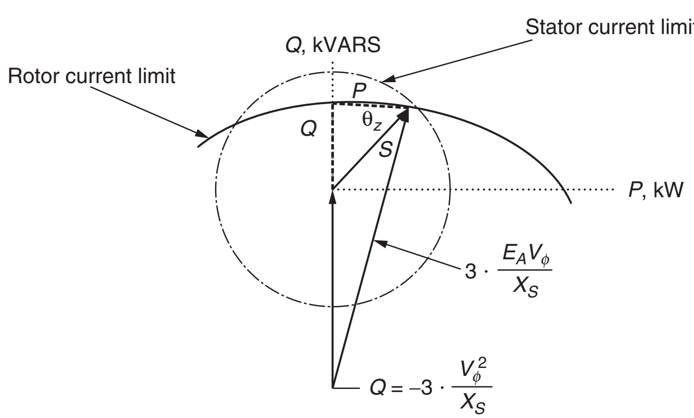


Note: $I_A \cos(\theta_z) = \dfrac{E_A \cdot \sin\delta}{X_S}$

(*a*) Developing the capability curve from the phasor diagram.



(*b*) Relating Power and Reactive Power to the phasor diagram.



(c) Relating the phasor diagram to the capability curve.

FIGURE 7-15 Estimating the real and reactive power of a generator as a function of field current and torque angle.

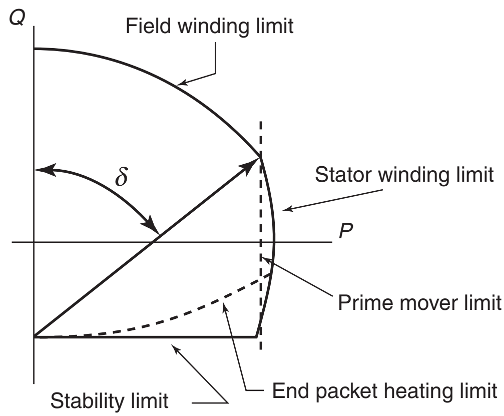


FIGURE 7-16 Capability diagram, round-rotor machine.

requirements. When these more sophisticated methods are available, it is generally better to use them. The approximate method described here is included for two reasons: (1) because it may aid in a physical understanding of generator operations and (2) because it gives reasonably accurate answers without the need for a large number of computer "cycles."

Figure 7-17 shows open- and short-circuit characteristic curves for an ac generator. These curves are taken with the generator operating at rated speed. Important curves are

1. The air gap line is the extrapolation of open-circuit voltage versus field current at low levels of field current.
2. The no-load saturation curve is the actual open-circuit voltage versus field current characteristic.
3. The short-circuit saturation curve typically does not show saturation. It represents a measurement of current in the terminals of the generator, with the terminals short-circuited, versus field current.
4. The rated current, zero power factor saturation curve shows voltage at the terminals of the machine versus field current if the machine is operated with rated armature current at zero power factor.

Calculating Excitation Requirement. Field current required for machine operation consists of three parts

$$I_f = I_{FG} + I_{FSI} + I_{FS} \tag{7-47}$$

where, I_{FG} is the field current required to excite the air gap, I_{FSI} is the field current required to compensate for direct-axis armature current, and I_{FS} is the field current required to compensate for saturation. Note that this method does not account for saturation of the quadrature axis. Armature resistance is generally neglected.

Addition of the three components of current is shown in Fig. 7-18. In this figure, the field current that compensates for armature reaction is added at the power factor angle Θ. The current I_{FS} is the current required to compensate for direct axis saturation. I_{FS} is the distance from the air-gap line to the no-load saturation curve at the voltage corresponding to the flux level in the magnetic circuit. This voltage is referred to as the voltage behind Potier reactance. It is estimated as shown in Fig. 7-19.

To find the (fictitious) Potier reactance, refer back to Fig. 7-17. Note that for the zero power factor test, all the fluxes are on the direct axis, so they add directly. Two of the current components, I_{FG} and I_{FSI}, are easily determined. The difference between field current for the zero power factor test and the sum of these two currents is I_{FS}. Since this corresponds to the distance between the air-gap

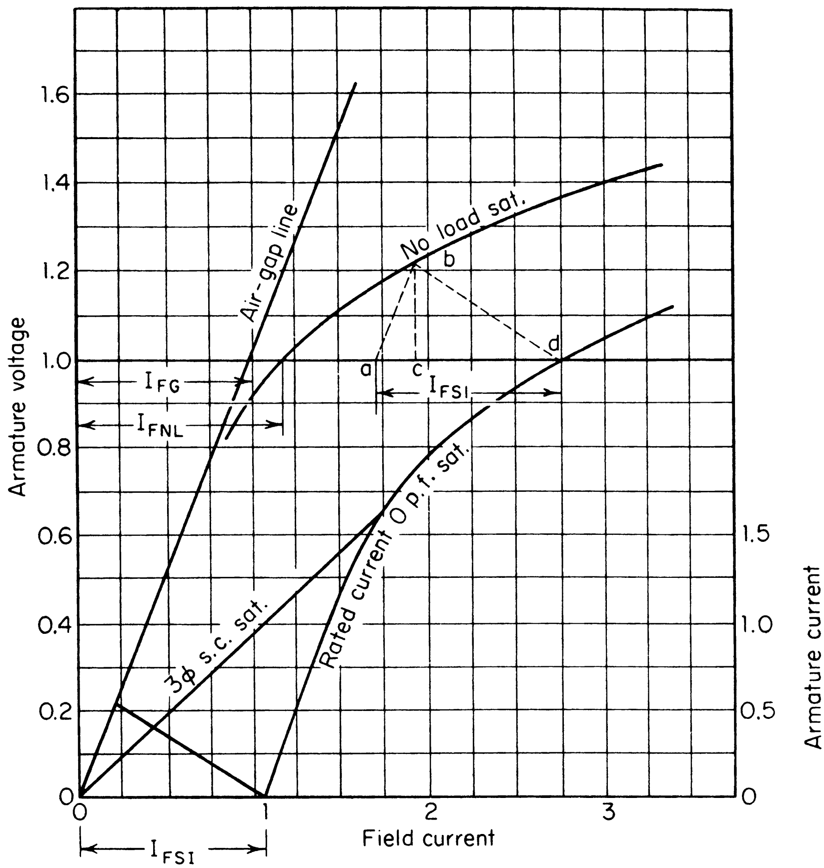


FIGURE 7-17 Typical saturation curves of an ac generator showing graphic determination of Potier reactance (quantities are in per unit values).

line and saturation curve at the voltage behind Potier reactance, it determines that voltage. For the zero power factor test, $I_a x_p$ is a vertical line of length x_p since $I_a = 1$. The distance between the saturation curve and air-gap line is the same as I_{FS} at a voltage found by casting a line parallel to the air-gap line from $I_F - I_{FSI}$. This is *ab* on Fig. 7-17. Then x_p corresponds to *bc* on the same figure.

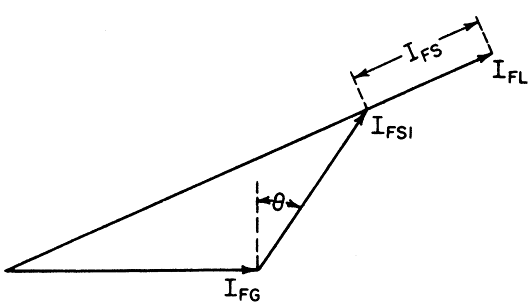


FIGURE 7-18 ANSI method of calculating load excitation.

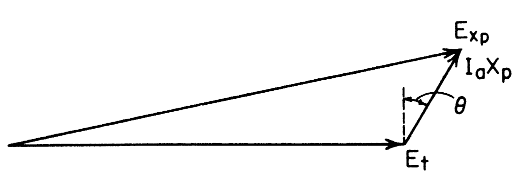


FIGURE 7-19 Calculation of voltage behind Potier reactance.

7.5 ARMATURE WINDINGS

A wide variety of winding types may be used to produce a desired voltage with the desired number of phases and a suitable waveshape. In small generators, "scramble wound" armature windings may be used. However, in most alternator applications, double-layer, form-wound coils in open slots with 60° phase belts are used. In such a winding, each slot has two conductor bars (often called *half-coils*), not necessarily from the same phase winding. These bars are insulated from ground and secured in the slot, usually by wedges. It is usually necessary for the bar to have the ability to slide axially in the slot to accommodate thermal expansion, but it must not be loose in either the radial or azimuthal directions. This has led to a number of proprietary techniques for armature construction.

7.5.1 Winding Forms

Figure 7-20 shows an example winding diagram. For the purposes of this figure, the machine is shown "rolled out flat," with the dotted lines on either side representing the same azimuthal location. In this case, the machine has 24 slots, each with two half-coils, as shown in the slot allocation section of the drawing, at the bottom of the figure. The upper part of the figure shows how one phase of the winding would be laid out. This drawing shows a lap type winding (the most commonly used) with a $^5/_6$ pitch. In a 24-slot, 2-pole winding a full-pitch coil would span 12 slots, while in the $^5/_6$ pitch winding the coils span 10 slots.

Fractional Slot Windings. Fractional slot windings, in which the number of slots per pole per phase is not an integer, have coil groups that differ from one another. These can be arranged to produce balanced voltages under circumstances that are beyond the scope of this discussion.

7.5.2 Stranding and Transposition

At power frequencies (50 or 60 Hz), the skin depth in copper is on the order of 1 cm so that it is usually necessary to subdivide armature conductors into a number of parallel strands. In form-wound

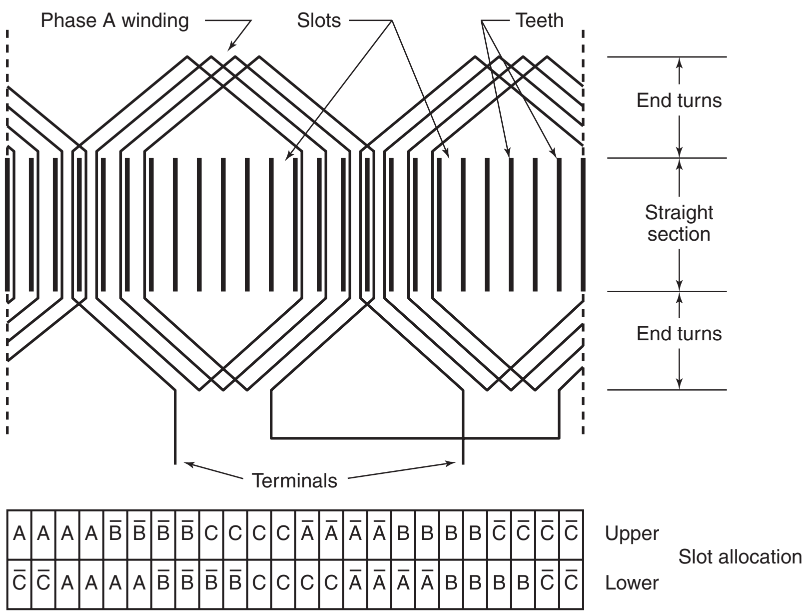


FIGURE 7-20 Armature in 24 slots, 5/16 pitch.

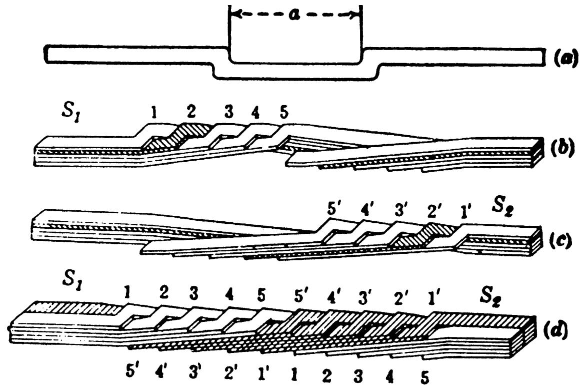


FIGURE 7-21 Illustration of Roebel transposition: (*a*) typical offset conductor strand; (*b*) group of conductor strands composing half the conductor; (*c*) complementary group; (*d*) assembly.

coils, these strands are usually rectangular to allow for good space factor. To prevent circulating currents between parallel strands, it is necessary to employ transposition to ensure that voltages induced in each strand are approximately the same.

The simplest form of transposition, often used in transformers and sometimes in generators, is to twist the armature conductors at 180° in the end turns. Or sometimes, groups of conductors are connected together in the end turns with a progressive transposition that constitutes a "twist" of the winding from half-coil to half-coil.

Transposition of strands in the end turns is generally not satisfactory in large ac generators. A transposition scheme attributed to Roebel is usually used (see Fig. 7-21). The Roebel transposition is equivalent to a twist of the conductors in the slot. It is usual to carry out the Roebel transposition only within the slot part of the winding. A variety of transpositions are used, including 180°, 360°, and 540°. The first two are effective only at eliminating circulating currents due to flux crossing the slot, but do not compensate for flux in the end windings. The 540° transposition, attributed to Ringland and Rosenberg (1959) is often applied because it filters out most of the circulating currents in a bar. Other more complex Roebel transposition arrangements that more extensively filter out circulating currents are possible but are rarely used because of manufacturing complexity. Bapat (1973) and Neidhoeffer (1990) outline general surveys about special transpositions.

7.6 INSULATION SYSTEMS

Electrical insulation is used to isolate field conductors from each other and ground, and in the armature winding to isolate strands and turns from each other and the whole winding from ground. Proper application of electrical insulation constitutes much of the art of ac machine design, particularly in the larger generator sizes.

7.6.1 Materials

International standards have established various classes of insulation systems based primarily on the maximum steady-state operating temperature and have established voltage proof tests to demonstrate the dielectric capability of insulation systems.

Of the several thermal classes of insulation systems, three are most applicable to ac generators. These are Classes 130, 155, and 180. Class 105 insulation materials are no longer used in modern ac generators. In the most recent revisions of IEEE C50-12 and C50-13, and IEC 60034-1 and 60034-3 standards for ac generators, the numerical classifications 130, 155, and 180 have been substituted for the previously used alphabetical classifications B, F, and H, respectively.

Class 130 systems comprise inorganic materials such as mica and glass fibers, organic materials such as synthetic films, and suitable binders. Until recently, Class 130 systems have been identified as Class B systems.

Class 155 systems comprise generally similar materials with binders capable of higher temperatures. Until recently, Class 155 systems have been identified as Class F systems.

Class 180 systems include silicone elastomers as well as mica, glass fibers, and higher-temperature binders. Until recently, Class 180 systems have been identified as Class H systems.

Synthetic chemistry has presented a continued flow of materials suitable for electrical insulation, and any of the above-mentioned insulation systems may be supplemented by new materials.

7.6.2 Temperature Measurements

Several methods of observing winding temperatures are used. Note that it is ordinarily not practical to actually measure the "hot spot" temperature directly, so that an allowance is normally given, based on the method of temperature measurement used.

Thermometer method. A thermometer (mercury filled, or other suitable temperature measuring device) may be applied to what is expected to be the hottest part accessible. Of the several methods described, this one gives readings furthest from the maximum winding temperature and therefore has the highest hotspot allowance.

Resistance method. The resistance of electrical conductors is a function of temperature. Thus, if the resistance of the winding can be measured and its resistance at a reference temperature is known, the average temperature of the winding may be calculated

$$T_2 = K + \frac{R_2}{R_1} \cdot (1 + T_1) \tag{7-48}$$

where R_1 is the resistance at temperature T_1 and R_2 is the resistance at temperature T_2. The constant K is 234.5°C for copper, 242°C for silver-alloyed copper used for hollow conductors, and 225°C for electrical conductivity aluminum.

Discharge coolant method. For directly cooled armature windings in the largest generators, the observable temperature is the coolant being discharged from the armature coils.

Embedded detector method. Thermocouples or resistive thermal detectors are built into the machine. In the most common application these are placed between the coil sides in a two-layer armature winding near the axial center or warmest region of the core.

Applied thermocouple method. Thermocouples are applied directly to the conductors, with little or no insulation. This method is used only for experimental testing to, for example, determine hot-spot allowances. It is not used for operational measurements.

7.6.3 Temperature Ratings

Temperature ratings of ac generator windings are based on measurements of winding temperatures and the thermal class of insulation used. These measurements are specified in IEEE C50-12 and C50-13, and IEC 60034-1. Generally, machines with smaller ratings employ the resistance method and larger machines employ embedded detectors or discharge coolant measurement methods.

7.6.4 Armature-Winding Insulation

Armature voltages range from about 220 V to about 27 kV. With such a wide range, different techniques are employed. In the armature, insulation is for strands, turns, and ground wall.

Strand insulation is required to prevent circulating currents within a conductor bar. The voltage levels are not high so mechanical integrity is the important feature of strand insulation. This is usually a layer of served fabric or film coating.

Turn insulation is used in multiturn coils, generally applicable only in small-size generators. This insulation is required to withstand turn-turn voltage, although in some cases large transient spikes of voltage may be incident on the winding.

Ground wall insulation must withstand full voltage to ground. Typically, the whole of an armature winding is insulated for full voltage, even though some of the coils, located near the neutral end of the winding, see lesser voltage.

In high-voltage armatures (above ~5 kV), some measures must be taken to control the effects of corona and partial discharge. In the slot portion of the coil, it is necessary to prevent discharges due to capacitive coupling through the insulation, from the surface of the insulation to the grounded stator core. These discharges are prevented by coating the outer surface of the insulated conductor with a conductive (sometimes called semiconducting) coating (paint or tape). To prevent discharges along the surface of the conductors in the end windings, those sections are sometimes coated with very weakly conducting coatings that are called grading coatings (paint of tape).

It is important to prevent electrical discharges in the vicinity of the winding because such discharges through air and in the presence of any water vapor will produce nitrous and nitric acid and ozone, substances corrosive to the materials of the winding.

7.6.5 Field-Winding Insulation

Field windings operate at much lower voltages (usually less than ~800 V). Some transient conditions, such as interruption of field current, can lead to much higher voltages.

Field windings are subject to the centrifugal forces due to rotation, and this presents special challenges. Dimensional stability is required of the field winding to prevent dynamic rotor imbalance. It is also necessary, in larger machines, to allow the field winding to expand thermally with respect to the rotor steel. The resulting "creepage" surfaces must allow slip in the axial direction but not movement in the other directions.

7.6.6 Insulation Maintenance

Failure of the insulation system of a generator produces a prompt outage of the generator. This produces ample reason to maintain the insulation system in good order. Insulation maintenance involves measures to reduce or eliminate sources of abnormal aging and damage to insulation, inspection to determine if deterioration has taken place, and testing to assess the state of the insulation system.

Service Conditions. Some operational conditions that produce accelerated insulation deterioration and ways of controlling them are described as follows:

Excessive temperature may be caused by blockage of ventilation paths or cooling tubes or by overload. Measurement of winding temperature is used to detect overtemperature conditions to allow the cause to be corrected.

Excessive voltage may be caused by control errors, but is usually the result of switching operations or lightning. This is usually a matter for correct system design to prevent such overvoltages.

Contamination can affect windings in one of the three ways. Some contaminants may attack insulation system, reducing its dielectric strength or affecting its mechanical properties. Some contaminants may thermally insulate parts, causing excessive temperatures. On the other hand, some contaminants such as conducting particles may lead to abnormal potential distribution and excessive voltage stresses. Control of contamination consists of preventing substances from entering the generator (if possible) and of inspecting periodically for signs of contamination or related damage.

Physical damage to insulation can occur for a variety of reasons, including vibration, overspeed, short-circuit forces or improper synchronization, and even damage by foreign objects. Generally, prevention and inspection are used to control such damage.

Corona, an electrical discharge around the surface of stator insulation, may be caused by contamination or by some inherent weakness in the insulation system itself.

Inspection. Visual inspection supplemented with simple physical tests may be carried out to diagnose incipient problems that might lead to eventual insulation failure. Some symptoms are described as follows:

Swelling or shrinking is evident from visual inspection, probing, and tapping on the insulation. Swelling is accompanied by reduction in physical integrity and sometimes by separation of insulation from the underlying conductor. Shrinkage leads to looseness in the slot.

Cracking results from abnormal mechanical stress or poor winding support. *Girth cracking* is displacement of insulation near the core ends of relatively large machines, caused by thermal cycling.

Substances noted on the winding can include contaminants such as oil, dust, and water and also deposits produced by corona, which may be white, gray, or red and may appear in regions of high electrical stress.

Slot wedges and fillers may become loose in service. This will lead to deterioration if not corrected, so any inspection of a generator should include checking of all the stator slot wedges.

Tests of Insulation Systems. Before commissioning a new generator and as part of generator maintenance programs, tests of insulation systems are carried out. There is a wealth of literature on this topic cited in the references. What appears here is only a first-order synopsis.

Testing of insulation systems in ac generators is used to uncover existing weaknesses or faults and to produce an indication of expected service reliability.

Some tests commonly employed are described as follows:

Insulation resistance is a straightforward test used to indicate the existence of contamination or moisture absorption. It seldom indicates the existence of an incipient fault. It is difficult to assign specific acceptable values for the results of insulation resistance tests, but it is possible to 'trend' the results of such tests.

Dielectric absorption tests are similar to a dc resistance test. A voltage in the range of 500 V to 5 kV is placed across the insulation and current is measured as a function of time. The *apparent* resistance is the ratio of voltage to current and is a function of time. The *polarization index* is the ratio of the value of this apparent resistance after 10 to the value after 1 min. The polarization index may be used for trending. Generally, for Class 130 and Class 155 insulation, a polarization index of about 2 or greater is an indication of a healthy winding, while a value of 1 or less is an indicator of potential trouble. Further description of resistance and dielectric absorption tests is contained in IEEE Standard 43.

Overvoltage tests are proof tests. Failure of an overvoltage test (called also *high potential* or *hi-pot*) is destructive. Such tests are carried out because failure of an insulation system under test conditions is much less damaging than failure in service. Overvoltage tests may be carried out with either ac or dc voltage, and employ voltage levels that are selected for the particular test. New windings (whether in a new generator or in a newly rewound generator) are typically tested to higher levels relative to their operating voltage, than are older windings that have been in service for a significant portion of their expected life. Typical values of high potential for testing new windings are twice operating line to line voltage plus 1000 V. Typical values of high potential for testing mature windings are between 125% and 150% of operating line to line voltage.

Direct current test voltage is related to the peak of the equivalent ac voltage. Details and test procedures are included in IEEE Standards 4 and 95.

Slot discharge may be detected in operation by using a *corona probe*, an antenna located inside the machine (Kurt et. al. 1984).

Rotor winding impedance tests can be used to detect failure of field-winding resistance. An alternating voltage is placed across the field winding, and individual coil voltages are measured. If the measured voltages are not the same, there is an indication of short-circuited turns.

Insulation power factor tests are generally used only on high-voltage machinery (above ~6 kV). Normally, the power factor, measured when applying voltage to a winding and measuring current, is quite low. Power factor measured as a function of applied voltage may be used for trending, and an increase in power factor with voltage is an indication of developing voids in the insulation.

Controlled overvoltage or dc leakage tests consist of measuring leakage current versus time. These tests are detailed in IEEE Standard 95.

7.6.7 Stator-Core Insulation

The stator core is made up of thin sheets of electrical steel that are coated by thin films of organic or inorganic insulation.

7.7 MECHANICAL CONSTRUCTION

Generators consist of a rotor, a stator and those mechanical parts required to hold the machine together and assure proper alignment, cooling, and other aspects of operation. Design criteria for the mechanical aspects of machine construction are based not only on anticipated steady state operation but also on abnormal aspects of operation such as overspeed and short circuits.

Essential variations in machine construction are salient pole versus round rotor and horizontal versus vertical shaft. In virtually all modern alternators the field is the rotating element and the armature is the stationary element. In salient-pole construction the field windings are wound around pole pieces. Such construction is used for machines with at least 6 poles, corresponding to shaft speeds of 1200 r/min or less in the 60-Hz world, or 1000 r/min in the 50-Hz world. Round-rotor construction, in which field windings are inserted in axial slots in a cylindrical rotor body, is used in essentially all 2-pole machines and most 4-pole machines because attaching salient poles to a rotor shaft becomes problematical at the higher shaft speeds associated with such machines.

The prime mover determines if the shaft is horizontal or vertical. Most generators driven by steam turbines, gas turbines, or reciprocating engines have horizontal shafts. Larger, low-speed hydraulic turbines have vertical shafts, while smaller, higher-speed turbines can have shafts that are vertical, horizontal, or even inclined. Magnetic circuits, stator, and armature windings of all machines are similar, only structural details differ.

7.7.1 Stator Construction

Armature cores are built up of thin laminations, produced as either segments or complete rings, depending on the size of the generator. Successive layers or groups of layers of the segmented laminations are staggered to minimize the effect of the joints in the magnetic circuit. The core is clamped between pressure plates and fingers to support it with sufficient pressure to prevent undue vibration of the laminations. Especially in long cores, the clamping arrangement may include some provision to compensate for compacting of the core after initial assembly.

The armature windings are fitted tightly in the slots and secured radially by wedges driven into suitable notches at the air gap end of the slots. It is necessary that the stator coil ends be able to resist the abnormal forces associated with short circuits. A supporting structure may be employed for this purpose. There are many variations of support design; most of them provide filler blocks between the coil sides, strategically located to transmit the circumferential forces from coil to coil, and additional structure to counteract the radial forces.

Coil supports ordinarily are designed to suit the need of a particular machine. Large 2-pole machines require a quite elaborate structure; the combination of large short-circuit currents and coil ends inherently flexible because of their long length makes these machines particularly susceptible to coil-end movement. Low-speed machines with stiffer coil ends require less support; in the smallest ratings the coils may be capable of withstanding the short-circuit sources without any additional support.

Stator frames, sometimes called casings, are commonly fabricated from structural steel, designed to support the core in alignment with the rotor and to suit the ventilating scheme used. In large machines with 2-pole or sometimes 4-pole construction, the stator core is mounted on springs to isolate core vibration from the machine frame.

7.7.2 Rotor Construction

The pole pieces of salient-pole alternators may be built up of steel laminations, both as manufacturing convenience and a means of limiting the loss in their air gap surfaces due to pulsations in air gap flux. The field coils, wound directly on the poles or preformed and then mounted on the poles, are suitably insulated from the poles for the voltages associated with normal and transient operation. The pole-and-coil assembly is bolted, dovetailed, or otherwise attached to the rotor body. It is the limitation of this attachment which usually dictates when round-rotor construction must be used rather than salient-pole construction.

The rotor body for a salient-pole machine may be a solid forging or assembly of heavy steel plates, for high-speed designs, or a spider-and-rim assembly for low-speed designs. The shaft may be integral with the body, as in the case of a forging, or may be bolted to or inserted into the body.

When the spider-and-rim construction is used, the entire assembly may be an integral weldment or casting, or the rim may be separate from the spider, as in the case of large waterwheel-driven generators. A common construction for this latter case is a rim built up of thin steel laminations, assembled around a cast or fabricated spider, bolted together between steel end plates and keyed to the spider.

The rotor of a round-rotor machine is cylindrical in shape with axial slots provided in its body for the field coils. The body is usually a steel forging with integral shaft ends. In special applications, other constructions may be used, with this same general configuration. The field coils are wound in axial slots in the rotor body, held in place by heavy slot wedges and by retaining rings over the coil ends.

Rotors are designed for operation at overspeeds, which depend on the characteristics of the prime mover. The overspeed limit (the speed above which the unit is no longer capable of safe operation) may be as low as 20% for a steam-turbine-driven unit or as high as 125% for some adjustable-blade, axial-flow hydraulic turbine-driven units.

7.7.3 Critical Speeds

The shaft system of the entire unit (generator plus prime mover) must be designed with regard to critical speeds. Both lateral and torsional critical speeds are considered. *Lateral critical speed* is the speed corresponding to the natural frequency of the shaft system in response to lateral or transverse forces such as residual unbalance forces. *Torsional critical speed* relates to the response of the shaft system to torsional forces. The lateral critical speed is associated with each mode of lateral vibration, the first critical speed corresponding to the lowest-frequency mode. Critical speeds are affected by shaft support, including foundation (particularly lateral critical speeds), and by internal and external damping.

It is preferable that the operating speed be at least 10% away from the nearest critical speed. Low-speed rotors ordinarily operate below the first critical speed. High-speed rotors, especially 2- and 4-pole turbogenerator rotors, usually operate above the first critical speed and sometimes above the second or even third critical speed. Multiplane balances are required for such rotors, and carefully planned speed profiles on start-up are required to avoid excessive stresses.

Torsional critical speeds are excited by external forces such as sudden load change, or a short circuit or other system disturbance, or cyclic variations in prime mover torque, such as from an internal combustion engine. Impulses from the buckets of a hydraulic turbine or from the gear teeth in a unit driven through the gear may also excite torsional vibrations.

It is possible that some components of a shaft system may have resonant frequencies, which are close to those that might be excited in operation. For example, long, last-stage turbine buckets might have a resonance at nearly twice the operating frequency, and that frequency is likely to be excited by negative-sequence currents. Care is required in selecting all components of the shaft system.

7.7.4 Bearings

Although, antifriction bearings are occasionally used on alternators of smaller ratings, the great majority are furnished with oil-lubricated babbitted bearings. For horizontal shafts at small ratings ring-oiled bearings are used, but at higher ratings recirculation of externally cooled oil is used.

Two principal types of thrust bearings are used on vertical alternators: the pivoted-shoe type and the spring type. The adjustable *pivoted-shoe* type, introduced in the United States by Albert Kingsbury, consists of a flat rotating collar or runner of steel or fine-grained cast iron resting on a stationary member consisting of several babbitted segmental shoes pivoted near their center on adjusting screws, which, by changing the elevation of the shoes, can provide equal loading on each. The screws also permit small adjustments in rotor elevation to correct generator and turbine clearances.

The bearings are immersed in oil. In operation, a thin, wedge-shaped film of oil is formed between the runner and the shoe. The oil is continuously circulated by the rotation of the runner and is cooled by either radiation or water cooling, usually within the oil bath but occasionally by an external system. Some of the larger bearings are cooled by means of water circulate through tubes embedded below the babbitt surface.

The *spring*-type bearing is inherently self-equalizing; that is, each shoe carries very nearly the same amount of load. A variation of the pivoted-shoe bearing, in which the shoes are supported on a system of interconnected levers, provides the same self-equalizing feature.

The *spherical bearing* is another variation of the pivoted-shoe thrust bearing, in which the runner is part of a sphere and the shoes of corresponding shape. This type of bearing restrains lateral movement of the shaft, serving the dual function of thrust and guide bearing.

Horizontal-shaft alternators occasionally require thrust bearing, as, for example, a single-impeller reaction turbine having unbalanced hydraulic thrust which must be restrained by the bearing. Thrust bearing designs for this application are generally of the pivoted-shoe type, either adjustable or equalizing.

Some thrust bearings, particularly of the adjustable pivoted-shoe type, may be provided with load cells for measuring and equalizing the thrust on the shoes. These may be of the hydraulic or strain-gage type, the latter is more common in modern applications. In addition to providing a check on the adjustment of the shoe loadings, these devices provide information about the hydraulic thrust characteristics of the turbine.

Guide bearings for vertical alternators are oil-lubricated babbitted rings. These are frequently segmented to facilitate assembly and may be composed of individual shoes which are radially adjustable. Guide bearings usually are partly immersed in an oil bath with oil circulated by the pumping action of sloping grooves in the babbitt surface. Occasionally, a separate lubrication system is provided which introduces oil at the top clearance of the bearing, collects it at the bottom, and recirculates it. It is common practice to place a guide bearing closely above the thrust bearing in the same oil pot. In some instances a guide bearing is on the outer periphery of the thrust runner.

Bearing Arrangements. Vertical alternators are classified by their bearing arrangement. In *suspended unit* the thrust bearing is located above the rotor and is provided with two guide bearings, one above the rotor and one below. The upper guide bearing is frequently placed just above and in the same oil pot as the thrust bearing. In "umbrella" unit, the thrust bearing is located below the rotor and one guide bearing also below the rotor, usually in the same oil pot and just above the thrust bearing. When umbrella arrangements are used, careful consideration must be given to the stability of the unit with respect to overturning moments. Large diameter, slow-speed units, in which the ratio of rotor diameter to the height of rotor center of gravity above the thrust surface is relatively large, lend themselves to umbrella construction. A principal advantage of umbrella construction is a substantial reduction in required powerhouse headroom, since the relatively high bearing support structure is below the rotor in the turbine pit. Furthermore, this structure itself usually spans a shorter distance when placed below the rotor and is not so high because of this, and shaft length may be reduced. A modified umbrella arrangement, having a guide bearing above the rotor, may be used when mechanical stability precludes classic umbrella construction. This retains some of the advantage of reduction in headroom requirement. The thrust bearing may be supported from the turbine head cover to achieve a more compact arrangement.

High-Pressure Lubrication. Horizontal journal bearings and thrust bearings are often provided with high-pressure oil ("jacking oil") to reduce starting torque and minimize bearing wear on start-up. Oil at pressures on the order of 10 MPa is introduced at the bottom of the journal bearing or at the center of each shoe or segment of a thrust bearing to lift the rotor and introduce an oil film in the bearing clearance before the shaft rotates.

Steam turbine generator units frequently are operated at very low speeds when they are disconnected from the system to prevent the shaft system from sagging while cooling or while at standstill for extended periods. This "turning gear" operation may be at 5 to about 50 r/min much below the speed at which the bearing would maintain an oil film. Jacking oil is usually used during this operation.

7.8 LOSSES AND EFFICIENCY

Efficiency of a generator is defined as

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} \cong 1 - \frac{P_{\text{dissipation}}}{P_{\text{out}}} \tag{7-49}$$

Windage and friction includes loss from the bearings, shearing of the air or hydrogen at the rotor surface, mechanical work done on the air or hydrogen that flows through the rotor, and power expended by the cooling fans. This loss is generally not a function of operating point, but it does vary with the pressure of the cooling gas inside the generator. Provided that the cooling gas pressure is kept constant, this loss is usually considered to be independent of load.

Core loss is caused by hysteresis and eddy currents in the core laminations. It also includes losses induced in structural parts of the machine that are exposed to stray alternating magnetic flux at no-load conditions. Core loss is a function only of terminal voltage, and it is usually considered to be independent of load.

Field loss is ohmic loss caused by field current flowing in the field winding. This loss depends on both the field current and the temperature of the field winding. It thus varies with load.

Armature loss is ohmic loss caused by armature current flowing in the armature winding. This loss is defined by the square of the armature current multiplied by the dc resistance of the armature winding, corrected for temperature. This loss varies with load.

Stray load loss explains sources of loss not adequately covered by the other categories. It includes eddy current loss in the armature and losses in structural elements exposed to magnetic fields arising from armature current and to those parts of the rotor surface affected by armature leakage fields. Stray loss is generally proportional to the square of armature current and may, with only small error, be expressed as a fraction of armature transport loss.

7.9 TESTING OF AC GENERATORS

Tests are performed on generators to establish conformance with projected performance and dynamic performance parameters. Details of such tests are contained in IEEE Standard 115, IEC 60034-2 and IEC 60034-4 standards.

7.9.1 Resistance

Field and armature resistances are typically small, so measurements should be made using a 4-wire technique. It is important that resistance be measured at a known temperature so that correction can be made to actual operating temperature.

7.9.2 Open-Circuit Saturation Curve

The generator is driven by a motor to rated speed and excitation varied to produce terminal voltage over a range, typically from perhaps 30% to 120% that of rated. Some caution is required here, particularly for large machines in which excessive flux can damage the core. Open-circuit losses may be established by this test if the drive motor is well characterized and input power is measured.

7.9.3 Short-Circuit Saturation Curve

This test is similar to the open-circuit test, except the armature terminals are short-circuited and excitation varied to produce armature current over some convenient range. Windage and friction losses may be inferred from power input at zero excitation. Stray load loss may be estimated as the difference between input power at rated armature current and the sum of friction and windage and armature I^2R.

7.9.4 Zero Power Factor Saturation Curve

For a relatively small generator, the zero power factor saturation curve can be determined by running the machine with its shaft unloaded, driven by a second generator.

By adjusting the excitation on the ac generator under test and excitation on the second generator, it is possible to measure the zero power factor saturation curve.

Rather extensive discussion of this method is described in IEEE 115. For large generators for which this "back-to-back" method is not practical, the zero power factor curve is usually determined by numerical methods. Often those methods employ finite elements.

7.9.5 Deceleration

Deceleration may be used for determining losses if the inertia of the machine is known. Since, if the shaft of a machine is unloaded, power dissipated is

$$P_w = \omega_m J \frac{d\omega_m}{dt} = \frac{d}{dt}\left\{\frac{1}{2} J\omega_m^2\right\} \tag{7-50}$$

where ω_m is mechanical speed, deceleration through synchronous speed can give a good measure of dissipation. The test may be run with the machine operating either at open-circuit or short-circuit conditions, or at zero excitation. It is usually run from a slight overspeed. This test can be used to determine an unknown inertia from known losses and observed deceleration.

7.9.6 Heat Runs

These are tests performed by operating the generator at some condition until the temperature stabilizes. Heat runs at open-circuit, short-circuit, and zero power factor may be combined to estimate temperature rise in actual operation. In large machines, good estimates of dissipation may be made by measuring the temperature rise of coolant (e.g., water). This is an alternative or supplement to measuring input power to the drive motor or machine deceleration. ω_m

7.10 COOLING

Dissipation in generators appears as heat which must be removed. This heat appears in the armature conductors, field-winding conductors, stator core, rotor surface, and other structural elements of the machine. Cooling of armature and field conductors may be direct or indirect; the difference is direct contact of the cooling medium with the conductor or contact through electrical insulation.

7.10.1 Cooling Media

Alternating-current generators may be cooled by air, hydrogen, water, or (very infrequently) oil. In large machines, no matter what the cooling medium, heat is transferred to water in heat exchangers that are located within the machine case.

Smaller machines are cooled by air. Recently, there has been a trend toward air-cooling larger machines. The upper limit in size for air-cooled machines is, as of this writing, about 350 MVA, and may increase further. The advantage of air cooling is simplicity. The disadvantage is machine size.

Hydrogen has had wide application in cooling of larger generators. It has a high specific heat and thermal conductivity and low density, so it provides better heat transfer with lower windage losses than does air. Hydrogen also does not support oxidation, with some advantage to insulation systems. Cooling a generator with hydrogen requires additional systems to maintain hydrogen purity and to remove hydrogen from lubricating oil and shaft seals. Since the "explosive" range of hydrogen/oxygen mixtures is about 5% to 75% hydrogen, if the purity of hydrogen is kept above about 95%, the cooling medium will be nonexplosive.

Water is used in armature winding cooling in very large machines.

7.10.2 Ventilation Paths

Fans used in electric machines may be of either radial flow or axial flow, and a wide variety of cooling paths are used. Figures 7-22 and 7-23 show two possible schemes.

7.10.3 Stator-Core Ventilation

The stator core is usually gas (air or hydrogen)-cooled. Axial passages through the core may be formed by punching holes in the laminations. Radial passages are formed by spacers that hold the core packets apart. Radial passages might be about 1 cm in axial length with spacing of about 5 cm.

In some cases, as shown in Fig. 7-22, axial and radial cooling passages are mixed in one machine. In some cases, the ventilating gas passes first radially inward and then radially outward.

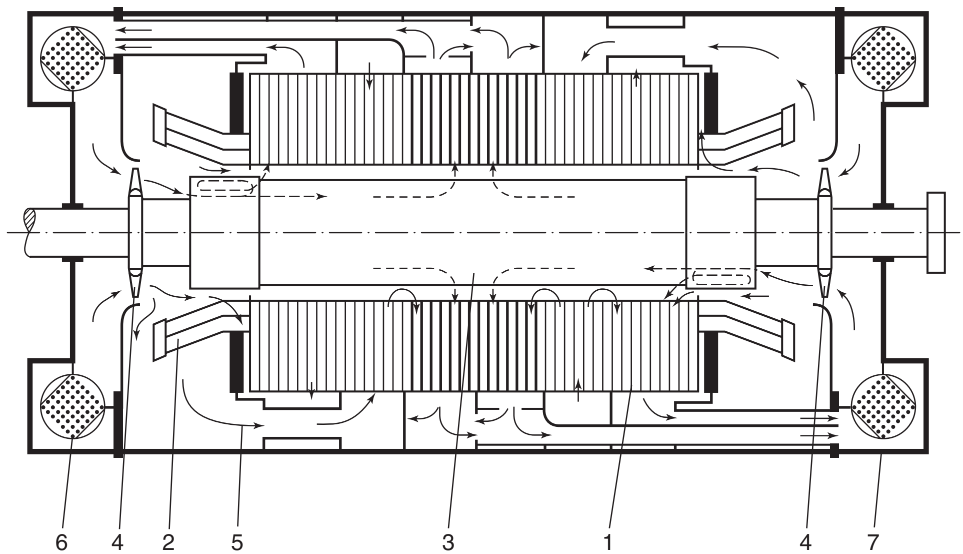


FIGURE 7-22 Cooling system, gas only. (1) stator core; (2) indirect-cooled stator winding; (3) rotor body; (4) axial-flow fans; (5) gas flow; (6) hydrogen cooler; (7) stator housing. (*ABB*.)

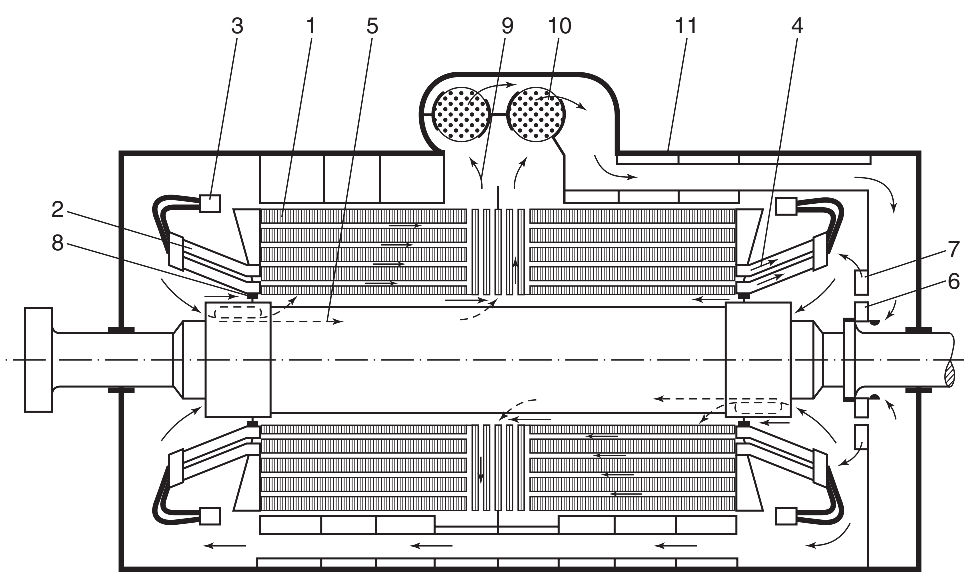


FIGURE 7-23 Cooling system, combination of gas and water. (1) stator core; (f2) water-cooled stator winding; (3) water manifold; (S45) water flow; (5) rotor body; (6) radial-flow fan; (7) diffuser; (8) air gap baffle; (9) gas flow; (10) hydrogen cooler; (11) stator housing. (*ABB*.)

7.10.4 Rotor Ventilation

As with the stator core, rotor cooling takes on a variety of forms. In some cases cooling gas passes axially from the ends of the rotor and then exits through holes in the rotor surface into the air gap, and then passes through the stator core. In other cases, gas passes radially inward from the air gap, diagonally through the rotor, and then radially outward to the air gap. This scheme can be coordinated with cooling of the stator core.

7.10.5 Direct and Indirect Cooling

Direct cooling, the norm for rotor windings and widely used in stator windings, exposes the cooling medium directly to the conductors. Figure 7-24 shows hydrogen and water directly cooled conductors for both stator and rotor.

In a directly gas-cooled stator, relatively large passages are built into the conductor bar. The conductor strands are transposed around the gas passages. There is strand insulation between the conductor strands and gas passage (which is often made of stainless steel), but the gas is within the ground wall. In a directly gas-cooled rotor the gas flow may be radial, axial, or diagonal, or some combination of all three.

In a directly water-cooled stator winding, the water flow may be in direct contact with the conductors. In some cases some or all of the conducting strands are made of hollow copper tubing. In others, stainless-steel tubes are used. Typically, water flows through the machine only one or two axial passes before being returned to the cooler. If water cooling is used, then

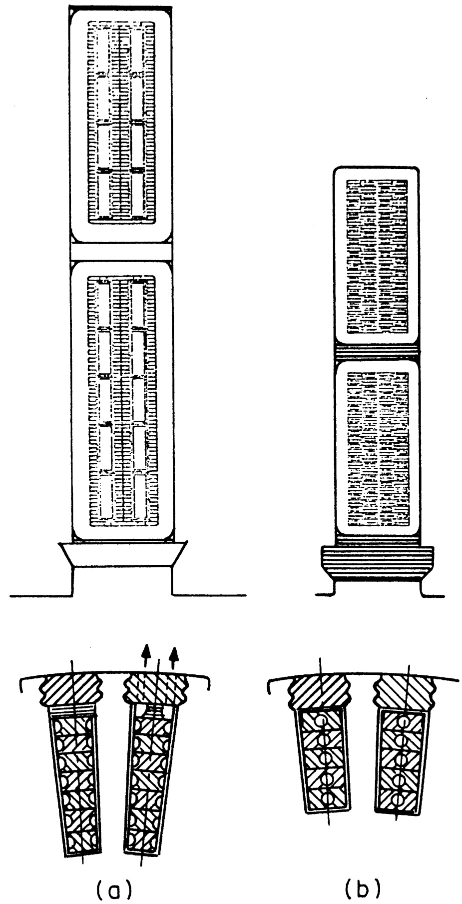


FIGURE 7-24 Directly cooled conductors; (*a*) hydrogen; (*b*) water.

1. The water is maintained at very high purity so that it has low conductivity.
2. Water is carried to the armature conductors through specially made hoses, since the conductor bars are at high potential and the water header is at ground.
3. Generally, hydrogen pressure in the machine is maintained above water pressure so that any leak will be of hydrogen into the water system, rather than water into the electrical insulation.

Water-cooled field windings are relatively rare, although many have been in highly reliable service for decades in some of the world's most powerful nuclear turbine generators.

7.11 DYNAMIC MODELS

In applying an ac generator to power systems, it is important to understand the dynamic performance of the machine. This section describes one way of estimating the dynamic performance of a generator. There is much literature on this topic, however, and what is done here will only skim the surface of a very deep topic.

Typical among the dynamic issues are levels of fault current, transient stability limits, responsiveness to excitation control, and damping of transient swings.

7.11.1 Per Unit System

Voltages, currents, and impedances are often expressed as *per unit* values, that is, a base for each quantity is assumed and every ordinary variable is compared with that base. Thus, when a machine is operating at its rated condition, both voltage and current magnitude could be said to be "one per unit." Multiplying voltage and current yields power, while dividing voltage by current yields impedance magnitude, so these can be expressed in per unit terms as well. The general form of the per unit (pu) is shown in Eq. 7-51.

$$pu = \frac{\text{actual}}{\text{base}} \tag{7-51}$$

Generally, the base apparent power or total power (S_{base}) and the base voltage (V_{base}) are specified; all other base values are determined from these two.

The use of per unit quantities has a number of advantages in analysis of electric machines. Not the least of these is that the use of the per unit index eliminates the need to refer to peak or rms, line neutral or line ground. In this development, all voltages, currents, fluxes, powers, and torques will be expressed in per unit. Time will not, however, be normalized, so rotational speed is measured in radians per second and time constants, in seconds.

7.11.2 Represented Circuits

Typical electrical models of electric machines represent the relationship between voltage and current in the various windings of the machine. Thus, the three phase windings and the field winding are represented. Important dynamics, however, arise from currents in elements of the rotor surface, the rotor body or amortisseur or damper bars, or some combination of all of these factors. These are typically represented by equivalent windings, too. For the purpose of this discussion, only one additional winding (referred to here as the "damper") will be included for each axis of the rotor.

The analysis begins with normalized variables referred to as the *direct* and *quadrature axes* of the machine. The per unit fluxes are

$$\begin{bmatrix} \psi_d \\ \psi_{kd} \\ \psi_f \end{bmatrix} = \begin{bmatrix} x_d & x_{ad} & x_{ad} \\ x_{ad} & x_{kd} & x_{fd} \\ x_{ad} & x_{fd} & x_f \end{bmatrix} \begin{bmatrix} i_d \\ i_{kd} \\ i_f \end{bmatrix} \tag{7-52}$$

$$\begin{bmatrix} \psi_d \\ \psi_{kd} \end{bmatrix} = \begin{bmatrix} x_q & x_{ad} \\ x_{aq} & x_{kd} \end{bmatrix} \begin{bmatrix} i_q \\ i_{kd} \end{bmatrix} \tag{7-53}$$

where the variables ψ represent per unit fluxes and i represents per unit currents. Subscripts d and q represent the direct and quadrature axes, respectively, while the subscripts a, k, and f represent the armature, damper, and field. There are only two armature variables here, where, strictly speaking, a third would be required. Ordinarily, the "zero" axis variable can be ignored as generators are usually connected so that no current can flow in that winding.

7.11.3 Equivalent Circuits

The equivalent circuits shown in Fig. 7-25 represent the same flux-current relationship as do Eqs. (7-52) and (7-53), with winding resistances added if

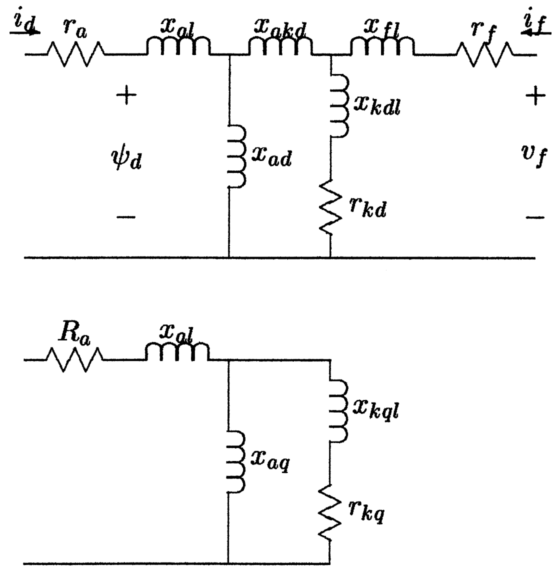


FIGURE 7-25 Direct and quadrature-axis equivalent circuits.

$$x_d = x_{al} + x_{ad} \tag{7-54}$$

$$x_{fd} = x_{ad} + x_{akd} \tag{7-55}$$

$$x_{kd} = x_{fd} + x_{kdl} \tag{7-56}$$

$$x_f = x_{fd} + x_{fl} \tag{7-57}$$

$$x_q = x_{al} + x_{aq} \tag{7-58}$$

$$x_{kq} = x_{kql} + x_{aq} \tag{7-59}$$

The interesting dynamics of the rotor can be now described by these two equivalent circuits.

7.11.4 Parameters

The various reactance and resistance parameters of the equivalent circuits of Fig. 7-25 may be calculated from first principles or may be measured by several different techniques. For example, IEEE Standard 115 describes a frequency response technique for measuring these parameters. Note that the parameter x_{al} appears in both direct- and quadrature axes. For this reason an apparently superfluous variable, x_{akd}, is used on the direct axis to allow an extra degree of freedom for an adequate "fit" between measurements and the equivalent circuit.

The equivalent circuits of Fig. 7-25 are only approximations to the actual performance of ac generators. In some cases it will be necessary to use higher-order models. Generally these are represented by multiples of the damper winding with different magnitudes and time constants. These are beyond the scope of this discussion.

7.11.5 Voltages

Voltages are produced by time variations of fluxes and by rotation of the machine. Translated into per unit measurement the voltages are

$$v_d = \frac{1}{\omega_0}\frac{d\psi_d}{dt} - \frac{\omega}{\omega_0}\psi_q + r_a i_d \tag{7-60}$$

$$v_q = \frac{\omega}{\omega_0}\psi_d + \frac{1}{\omega_0}\frac{d\psi_q}{dt} + r_a i_q \tag{7-61}$$

$$v_f = \frac{1}{\omega_0}\frac{d\omega_f}{dt} + r_f i_f \tag{7-62 }$$

$$v_{kd} = \frac{1}{\omega_0}\frac{d\psi_{kd}}{dt} + r_{kd} i_{kd} \tag{7-63}$$

$$v_{kq} = \frac{1}{\omega_0}\frac{d\psi_{kq}}{dt} + r_{kq} i_{kq} \tag{7-64}$$

$$v_0 = \frac{1}{\omega_0}\frac{d\psi_0}{dt} + r_a i_0 \tag{7-65}$$

7.11.6 Simulation Model

These expressions may be turned into a concise simulation model, suitable for use in modern computing apparatus for estimating performance of an ac generator. This is simply done by isolating the first-order time derivatives.

The state variables are the two stator fluxes ψ_d, ψ_q, two "damper" fluxes ψ_{kd}, ψ_{kq}, field ψ_f, and rotor speed ω and torque angle δ. The most straightforward way of stating the model employs currents as auxiliary variables

$$\begin{bmatrix} i_d \\ i_{kd} \\ i_f \end{bmatrix} = \begin{bmatrix} x_d & x_{ad} & x_{ad} \\ x_{ad} & x_{kd} & x_{fd} \\ x_{ad} & x_{fd} & x_f \end{bmatrix}^{-1} \begin{bmatrix} \psi_d \\ \psi_{kd} \\ \psi_f \end{bmatrix} \tag{7-66}$$

$$\begin{bmatrix} i_d \\ i_{kd} \end{bmatrix} = \begin{bmatrix} x_q & x_{ad} \\ x_{aq} & x_{kd} \end{bmatrix}^{-1} \begin{bmatrix} \psi_q \\ \psi_{kd} \end{bmatrix} \tag{7-67}$$

Then the state equations are

$$\frac{d\psi_d}{dt} = \omega_0 v_d + \omega\psi_q - \omega_0 r_a i_d \tag{7-68}$$

$$\frac{d\psi_q}{dt} = \omega_0 v_q - \omega\psi_d - \omega_0 r_a i_q \tag{7-69}$$

$$\frac{d\psi_{kd}}{dt} = -\omega_0 r_{kd} i_{kd} \tag{7-70}$$

$$\frac{d\psi_{kq}}{dt} = -\omega_0 r_{kq} i_{kq} \tag{7-71}$$

$$\frac{d\psi_f}{dt} = -\omega_0 r_f i_f \tag{7-72}$$

$$\frac{d\omega}{dt} = \frac{\omega_0}{2H}(T_e + T_m) \tag{7-73}$$

$$\frac{d\delta}{dt} = \omega - \omega_0 \tag{7-74}$$

and

$$T_e = \psi_d i_q - \psi_q i_d \tag{7-75}$$

For a practical simulation, v_d, v_q, and mechanical torque T_m must be specified.

7.11.7 Approximate Analysis

It should be clear from examining the equivalent circuits of Fig. 7-25 that the behavior of the machine, meaning its flux/current relationship, is a function of the speed of any given disturbance. This leads to approximate analyses of ac generators, which recognize the frequency dependence of the rotor elements.

Synchronous reactances x_d and x_q are applicable when the rotor is moving in synchronism with the magnetomotive force (mmf) produced by the armature currents, or when the deviation in speed is very small.

Transient reactance x'_d is applicable when armature mmf is changing with time, as in electromechanical transients or "swings." This effective reactance is reduced by the tendency of (nearly) short-circuited field winding to trap or hold flux nearly constant for periods comparable to or shorter than its time constant. This reactance is

$$x'_d \approx x_{al} + x_{ad} \,\|\, (x_{fd} + x_{fl}) \tag{7-76}$$

In some types of generators there may be an identifiable transient reactance on the quadrature axis, less than the quadrature-axis synchronous reactance, but in others the applicable reactance to use for the quadrature axis in transient events is just x_q.

For transient events that occur very rapidly, such as switching events and terminal short circuits, the effective reactances are the *subtransient* reactances. These result from the tendency of the amortisseurs (or even rotor iron) to support or trap flux. They are

$$x''_d = x_{al} + x_{ad} \,\|\, (x_{akd} + x_{kdl} \,\|\, x_{fl}) \qquad x''_d = x_{al} + x_{aq} \,|\, x_{kql} \tag{7-77}$$

The foregoing reactances are applicable when the armature mmf and rotor are rotating in synchronism or nearly so. The *negative-sequence* reactance x_2 is applicable with an armature mmf rotating backward at synchronous speed while the rotor is rotating forward. Negative-sequence currents arise from certain types of unbalanced operation. The negative-sequence reactance is approximately the average of subtransient reactances

$$x_2 = {}^1\!/_2(x''_d + x'_q) \tag{7-78}$$

Zero-sequence reactance is applicable to situations in which all 3 phases of the armature have identical currents such as would arise from a ground fault. Such currents do not produce substantial air-gap flux, and even some of the components of armature leakage are reduced, so the zero-sequence inductance is quite small.

7.11.8 Static and Transient Torque-Angle Curves

Torque-angle curves for operation of a synchronous generator may be written for both the synchronous and transient mode of operation. In per-unit these, torque-angle curves are for the steady state case

$$t_e = \frac{e_{af}}{x_d}\sin\delta + \left(\frac{1}{x_q} - \frac{1}{x_d}\right)\sin 2\delta \tag{7-79}$$

and for the transient case

$$t_e = \frac{e'_q}{x'_d}\sin\delta + \left(\frac{1}{x_q} - \frac{1}{x'_d}\right)\sin 2\delta \tag{7-80}$$

where δ is the torque angle, e_{af} is the internal voltage, and e'_q is the *voltage behind transient reactance*, and the terminal voltage is one per unit.

7.11.9 Stability by Equal Area

The relationship between power, which is directly proportional to torque, and displacement (power) angle is shown in Fig. 7-26. Note that the torque-angle curve for transient operation has a higher peak torque, although the two curves coincide at the steady-state operating point.

Disturbances from steady-state operation due to changes in prime mover torque, load changes, or faults cause load angle to change, usually with dynamic swings as described in this section. Generally, the most serious transient swing results from complete loss of load as from a nearby short-circuit. Under such a circumstance the generator, under the influence of prime mover torque, accelerates. The torque angle increases quadratically with time. Even after the power output is restored, the machine will swing forward until load torque stops its advance. A simple approximate way of estimating the maximum angle achieved is to use the *equal-area* criterion. The area of a disturbance is related to energy contributed to rotation, and for every (stable) situation the positive and negative areas must balance.

Pictured in Fig. 7-26 is a "critical" swing. The area underneath the prime mover torque curve from δ_0 to δ_c is the area that would be contributed if all load torque were zero during a period of acceleration between those two angles. If the area *above* the prime-mover torque between δ_c and δ_f is equal to or greater than this first area, the machine will regain synchronous operation. This establishes a *critical angle*. In turn, the *critical clearing time* t_c is established by

$$\delta_c - \delta_0 = \frac{1}{2}\frac{\omega_0}{2H}t_c^2 \tag{7-81}$$

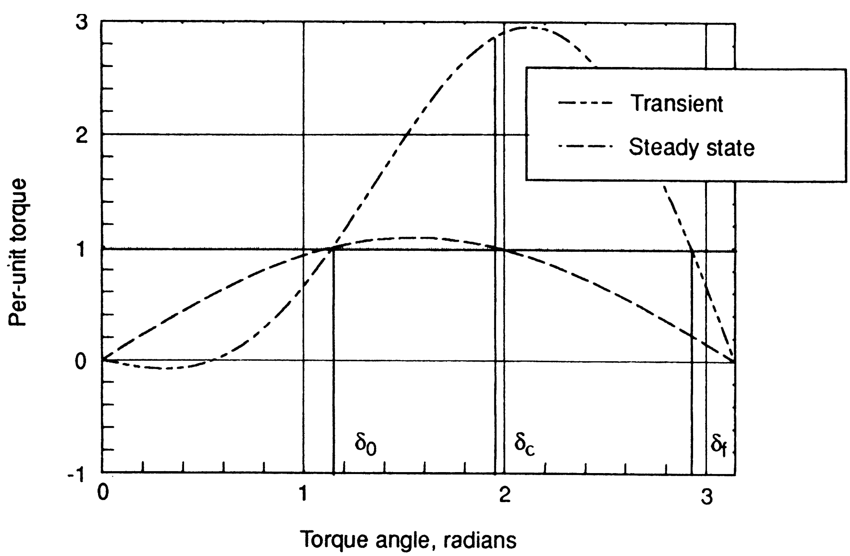


FIGURE 7-26 Torque-angle curves, steady state and transient.

7.11.10 Faults

If an ac generator is operated open-circuited at rated terminal voltage and then its terminals are suddenly short-circuited, the following currents flow in the terminal leads:

$$i_a = \left\{ \frac{e_{af}}{x_d} + \left(\frac{e_{af}}{x'_d} - \frac{e_{af}}{x_d}\right) e^{-(t/T'_d)} + \left(\frac{e_{af}}{x''_d} - \frac{e_{af}}{x''_d}\right) e^{-(t/T''_d)} \right\} \cos(\omega t - \theta_0) - \frac{e_{af}}{x''_d} \cos\theta_0 e^{-(t/Ta)}$$

$$i_b = \left\{ \frac{e_{af}}{x_d} + \left(\frac{e_{af}}{x'_d} - \frac{e_{af}}{x_d}\right) e^{-(t/T'_d)} + \left(\frac{e_{af}}{x''_d} - \frac{e_{af}}{x'_d}\right) e^{-(t/T''_d)} \right\} \cos\left(\omega t - \theta_0 - \frac{2\pi}{3}\right) - \frac{e_{af}}{x''_d} \cos\left(\theta_0 - \frac{2\pi}{3} e^{-(t/T_a)}\right)$$

$$i_c = \left\{ \frac{e_{af}}{x_d} + \left(\frac{e_{af}}{x'_d} - \frac{e_{af}}{x_d}\right) e^{-(t/T'_d)} + \left(\frac{e_{af}}{x''_d} - \frac{e_{af}}{x'_d}\right) e^{-(t/T''_d)} \right\} \cos\left(\omega t - \theta_0 + \frac{2\pi}{3}\right) - \frac{e_{af}}{x''_d} \cos\left(\theta_0 - \frac{2\pi}{3} e^{-(t/T_a)}\right)$$

These currents are shown in Fig. 7-27.

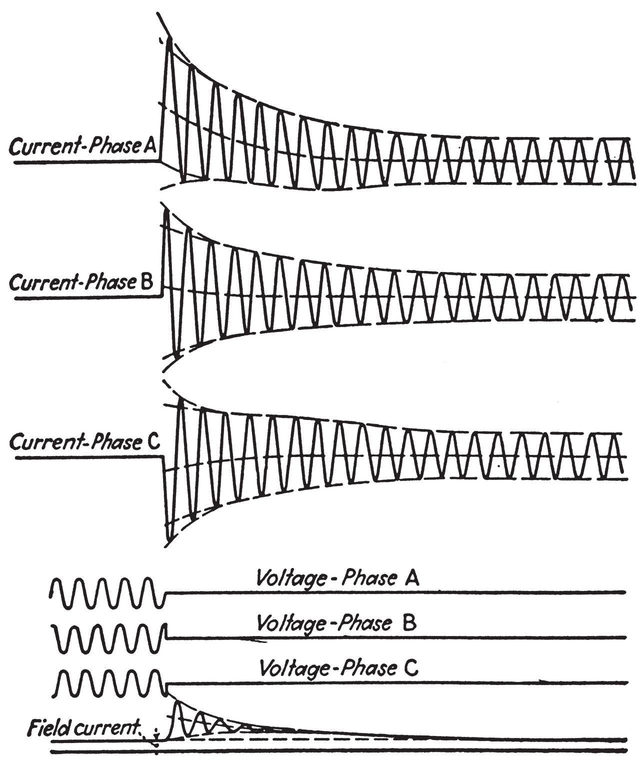


FIGURE 7-27 3-Phase (symmetric) fault currents.

BIBLIOGRAPHY

Textbooks

Chapman, S. J., *Electric Machinery Fundamentals*, 4th ed., New York: McGraw-Hill, 2005.

Fitzgerald, A. E., Kingsley, C., Jr., and Umans, S. D., *Electric Machinery*, New York: McGraw-Hill, 1990.

Hubert, C. I., *Electric Machines: Theory, Operating Applications, and Controls*, 2nd ed., Prentice Hall, 2002.

Industry Standards

ANSI C42.10, *Definitions of Electrical Terms.*

ANSI C50.22, *Recommended Guide for Testing Insulation Resistance of Rotating Machinery.*

IEC 60034-1, *Rotating Electrical Machines, Ratings and Performance.*

IEC 60034-2, *Rotating Electrical Machines—Methods for determining losses and efficiency from tests (excluding machines for traction vehicles).*

IEC 60034-3, *Rotating Electrical Machines—Specific requirements for cylindrical rotor synchronous machines.*

IEC 60034-4, *Rotating Electrical Machines—Methods for determining synchronous machine quantities from test.*

IEC 60034-18, *Functional evaluation of insulating systems (various parts).*

IEC 62114, *Electrical Insulation Systems (EIS)—Thermal Classification.*

IEEE C50.12, *Standard for Salient-Pole 50 and 60 Hz Synchronous Generators and Generator/Motors for Hydraulic Turbine Applications Rated 5 MVA and Above.*

IEEE C50.13, *Standard for Cylindrical-Rotor 50 and 60 Hz synchronous Generators Rated 10 MVA and Above*.

IEEE 115, *Test Procedures for Synchronous Machines.*

IEEE 1110, *Guide for Synchronous Generator Modelling Practices and Applications in Power System Stability Analyses.*

Technical Papers.

Bartheld, R., Organizational Structure of IEC TC2, *Conference Proceedings, IEEE Winter Power Meeting*, New York, 1999.

Berrong, D. B., McCown, W. R., Winnie, P. D., and Montgomery, L. W., Designing Central Station Turbine Generators for the Year 2000 and Beyond, *CIGRE Session*, 1998.

Detinko, F. M., Cooper, G. D., and Montgomery, L. W., Mechanical Design of a New Hydrogen Inner-Cooled Modular Generator Line, *International Joint Power Generation Conference and Exposition*, October 1992.

Drinkut, S. A., and Hurley, J. D., AC Generators and Generator Protection, Section 4.1, *Standard Handbook of Power Plant Engineering*, 2nd ed., New York: McGraw-Hill, 1997.

Emery, F. T., and Weddleton, R. F., Latest Advances Associated with Insulation Systems of High Voltage Stator Coils, *IEEE Symposium on Electrical Insulation*, June 1996.

Emery, F. T., The Application of Conductive and Stress Grading Tapes to Vacuum Pressure Impregnated, High Voltage Stator Coils, *IEEE Electrical Insulation Magazine*, July/August, 1996.

Gott, B. E. B., Advances in Turbogenerator Technology, *IEEE Electrical Insulation Magazine*, July/August, 1996.

Gott, B. E. B., McCown, W. R., Montgomery, L. W., and Michalec, J. R., Implications of Differences between the ANSI C50 Series and the IEC 60034 Series Standards for Large Cylindrical Rotor Synchronous Machines, *Panel Discussion IEEE-PES Summer Meeting, Berlin, Germany*, July 1997.

Gott, B. E. B., McCown, W. R., Montgomery, L. W., and Michalec, J. R., Implications of Differences between the ANSI C50 Series and the IEC34 Series Standards for Large Cylindrical Rotor Synchronous Machines, *IEEE PES Panel Session on Harmonizing.*

Gott, B. E. B., McCown, W. R., Montgomery, L. W., and Michalec, J. R., Update of Revision of ANSI C50 Series of Standards for Large Synchronous Machinery and Harmonization with IEC 34 Series, *Conference Proceedings, IEEE Winter Power Meeting*, New York, 1999.

Gott, B. E. B., McCown, W. R., Montgomery, L. W., and Michalec, J. R., Progress in Revision of IEEE/ANSI C50 Series of Standards for Large Steam and Combustion Turbine Generators and Harmonization with the IEC 34 Series, *International Conference on Electric Machines and Drives*, Boston, MA, June 2001.

Granicher, W., Air Cooled Turbogenerators up to 240 MVA, *Proceedings of the American Power Conference,* Vol. 52, 1990.

Haase, H., Largadier, H., and Suter, J., Air Cooled Turbogenerators in the 200 MVA Class, *Brown Boveri Review,* Vol. 73, March 1986.

Intichar, L., and Kulig, T. S., Development of Turbogenerators in Recent Years and in the Future, pp. 136–141, Vol. 1, *Proceedings of International Conference on the Evolution and Modern Aspects of Synchronous Machines,* August 1991.

Kaminski, C. A., Panel Discussion of Issues Related to Harmonization of Standards for Electrical Machines, *Panel Discussion, IEEE-PES Winter Meeting,* New York, 1999.

McCown, W. R., Winnie, P. D., and Montgomery, L. W., Trends in Electric Generator Development, *American Power Conference,* Vol. 59, 1997.

Nelson, R. J., and Montgomery, L. W., Electrical Design of a Modular Line of Two Pole Hydrogen Inner-Cooled Generators, *American Power Conference,* Vol. 54, 1992.

Nelson, R. J., Drinkut, S. A., and Gregory, M. D., ANSI vs. IEC Standards for Turbine Generators, *Proceedings of the American Power Conference,* Vol. 54, 1992.

Nilsson, N. E., Report on the Working Group to Revise ANSI C50.41, *Conference Proceedings, IEEE, Winter Power Meeting,* New York, 1999.

Nippes, P. I., and Nilsson, N. E., International Harmonization of Standards Detailed Report, *IEEE Transactions on Energy Conversion,* Vol. 14, No. 4, December 1999, pp. 1318–1322.

Nippes, P. I., and Nilsson, N. E., International Harmonization of Standards, *IEEE IEMDC'97,* 1997.

Nippes, P. I., IEC-US Standards Comparison by the IEEE-PES-EMC Task Force on Standards Harmonization, February 27, 1997.

Ruelle, G., Guillard, J. M., Bennett, R., and Jackson, R., Development of Large Air Cooled Generators for Gas Turbines and Combined Cycles, Paper 11-201, *CIGRE Session,* 1992.

Sedlazeck, K., Adelmann, W. et al; Influence of Customer's Specifications Upon Design Features of the EPR Turbogenerator, *CIGRE Session*, 2002.

Stephan, C., Baer, J., Zimmerman, H., Neidhofer, G., and Egli, R., New Air-cooled Turbogenerator in the 300 MVA Class, *ABB Review*, Jan. 1996.

Woods, E. J., Standards Harmonization: What Next, *Conference Proceedings of the IEEE Winter Power Meeting*, New York, 1999.

SECTION 8
DIRECT-CURRENT GENERATORS

O. A. Mohammed
Professor, Department of Electrical and Computer Engineering, Florida International University Miami, FL

CONTENTS

8.1 THE DC MACHINE

Applications. The most important role played by the dc generator is the power supply for the important dc motor. It supplies essentially ripple-free power and precisely held voltage at any desired value from zero to rated. This is truly dc power, and it permits the best possible commutation on the motor because it is free of the severe waveshapes of dc power from rectifiers. It has excellent response and is particularly suitable for precise output control by feedback control regulators. It is also well suited for supplying accurately controlled and responsive excitation power for both ac and dc machines.

The dc motor plays an ever-increasing vital part in modern industry, because it can operate at and maintain accurately any speed from zero to its top rating. For example, high-speed multistand steel mills for thin steel would not be possible without dc motors. Each stand must be held precisely at an exact speed which is higher than that of the preceding stand to suit the reduction in thickness of the steel in that stand and to maintain the proper tension in the steel between stands.

General Construction. Figure 8-1 shows the parts of a medium or large dc generator. All sizes differ from ac machines in having a commutator and the armature on the rotor. They also have salient poles on the stator, and, except for a few small ones, they have commutating poles between the main poles.

Former contributors include Thomas W. Nehl and E. H. Myers.

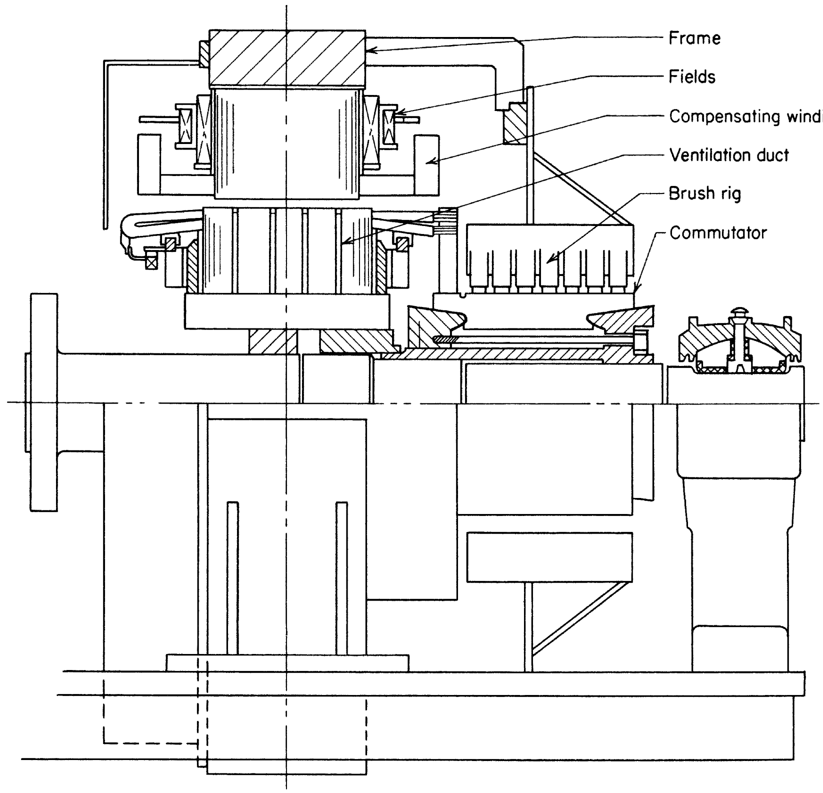


FIGURE 8-1 The dc machine.

Construction and Size. Small dc machines have large surface-to-volume ratios and short paths for heat to reach dissipating surfaces. Cooling requires little more than means to blow air over the rotor and between the poles. Rotor punchings are mounted solidly on the shaft, with no air passages through them.

Larger units, with longer, deeper cores, use the same construction, but with longitudinal holes through the core punchings for cooling air.

Medium and large machines must have large heat-dissipation surfaces and effectively placed cooling air, or "hot spots" will develop. Their core punchings are mounted on arms to permit large volumes of cool air to reach the many core ventilation ducts and also the ventilation spaces between the coil end extensions.

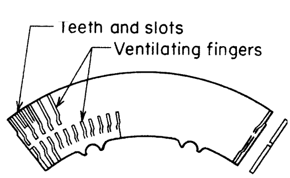


FIGURE 8-2 Armature segment for a dc generator showing vent fingers applied.

Design Components. Armature-core punchings are usually of high-permeability electrical sheet steel, 0.017 to 0.025 in thick, and have an insulating film between them. Small and medium units use "doughnut" circular punchings, but large units, above about 45 inches in diameter, use segmental punchings shaped as shown in Fig. 8-2, which also shows the fingers used to form the ventilating ducts.

Main- and commutating-pole punchings are usually thicker than rotor punchings because only the pole faces are subjected to high-frequency flux changes. These range from 0.062 to 0.125 in thick, and they are normally riveted.

The frame yoke is usually made from rolled mild steel plate, but, on high-demand large generators for rapidly changing loads, laminations may be used. The solid frame has a magnetic time constant of ½ s or more, depending on the frame thickness. The laminated frame ranges from 0.05 to 0.005 s.

The commutator is truly the heart of the dc machine. It must operate with temperature variations of at least 55°C and with peripheral speeds that may reach 7000 ft/min. Yet it must remain smooth concentrically within 0.002 to 0.003 in and true, bar to bar, within about 0.0001 in.

The commutator is made up of hard copper bars drawn accurately in a wedge shape. These are separated from each other by mica plate segments, whose thicknesses must be held accurately for nearly perfect indexing of the bars and for no skew. This thickness is 0.020 to 0.050 in, depending on the size of the generator and on the maximum voltage that can be expected between bars during operation. The mica segments and bars are clamped between two metal V-rings and insulated from them by cones of mica. On very high speed commutators of about 10,000 ft/min, shrink rings of steel are used to hold the bars. Mica is used under the rings.

Carbon brushes ride on the commutator bars and carry the load current from the rotor coils to the external circuit. The brush holders hold the brushes against the commutator surface by springs to maintain a fairly constant pressure and smooth riding.

8.2 GENERAL PRINCIPLES

Electromagnetic Induction. A magnetic field is represented by continuous lines of flux considered to emerge from a north pole and to enter a south pole. When the number of such lines linked by a coil is changed (Fig. 8-3), a voltage is induced in the coil equal to 1 V for a change of 10^8 linkages/s (Mx/s) for each turn of the coil, or $E = (\Delta\phi T \times 10^{-8})/t$ V.

If the flux lines are deformed by the motion of the coil conductor before they are broken, the direction of the induced voltage is considered to be into the conductor if the arrows for the distorted flux are shown to be pointing clockwise and outward if counterclockwise. This is generator action (Fig. 8-4).

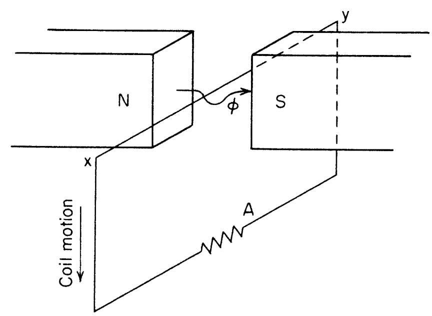


FIGURE 8-3 Generated emf by coil movement in a magnetic field.

Force on Current-Carrying Conductors in a Magnetic Field. If a conductor carries current, loops of flux are produced around it (Fig. 8-5). The direction of the flux is clockwise if the current flows away from the viewer into the conductor, and counterclockwise if the current in the conductor flows toward the viewer.

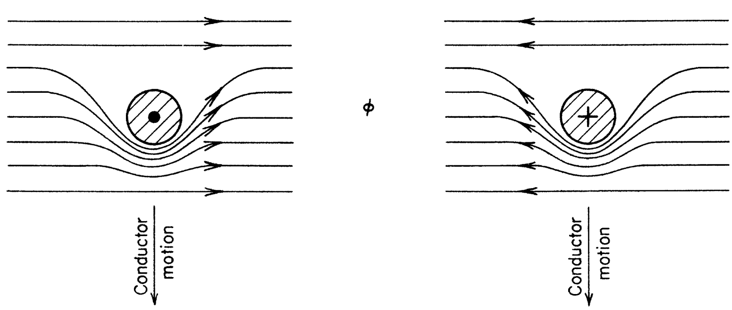


FIGURE 8-4 Direction of induced emf by conductor movement in a magnetic field.

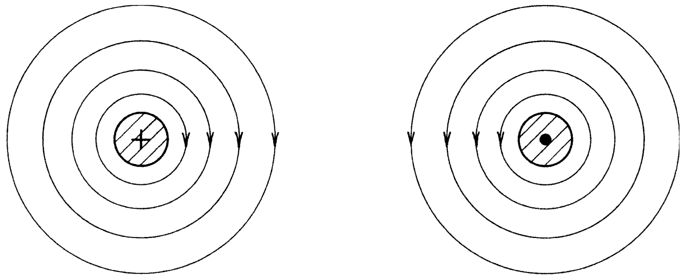


FIGURE 8-5 Magnetic fields caused by current-carrying conductors.

If this conductor is in a magnetic field, the combination of the flux of the field and the flux produced by the conductor may be considered to cause a flux concentration on the side of the conductor, where the two fluxes are additive and a diminution on the side where they oppose. A force on the conductor results that tends to move it toward the side with reduced flux (Fig. 8-6). This is motor action.

Generator and Motor Reactions. It is evident that a dc generator will have its useful voltage induced by the reactions described above, and an external driving means must be supplied to rotate the armature so that the conductor loops will move through the flux lines from the stationary poles. However, these conductors must carry current for the generator to be useful, and this will cause retarding forces on them. The prime mover must overcome these forces.

In the case of the dc motor, the conductor loops will move through the flux, and voltages will be induced in them. These induced voltages are called the "counter emf," and they oppose the flow of currents which produce the forces that rotate the armature. Therefore, this emf must be overcome by an excess voltage applied to the coils by the external voltage source.

Direct-Current Features. Direct-current machines require many conductors and two or more stationary flux-producing poles to provide the needed generated voltage or the necessary torque. The direction of current flow in the armature conductors under each particular pole must always be correct for the desired results (Fig. 8-7). Therefore, the current in the conductors must reverse at some time while the conductors pass through the space between adjacent north and south poles.

This is accomplished by carbon brushes connected to the external circuit. The brushes make contact with the conductors by means of the commutator.

To describe commutation, the Gramme-ring armature winding (which is not used in actual machines) is shown in Fig. 8-8. All the conductors are connected in series and are wound around a steel ring. The ring provides a path for the flux from the north to the south pole. Note that only the outer portions of the conductors cut the flux as the ring rotates. Voltages are induced as shown. With no external circuit, no currents flow, because the voltages induced in the two halves are in opposition.

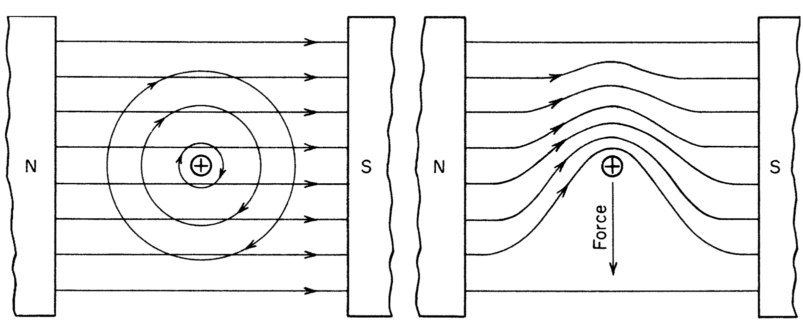


FIGURE 8-6 Force on a current-carrying conductor in a magnetic field.

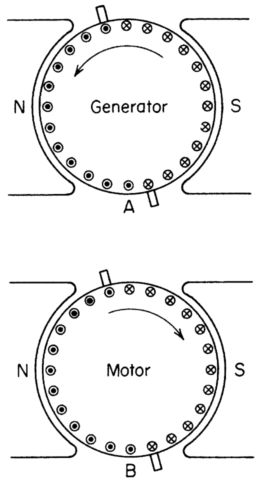


FIGURE 8-7 Direction of current in generator and motor.

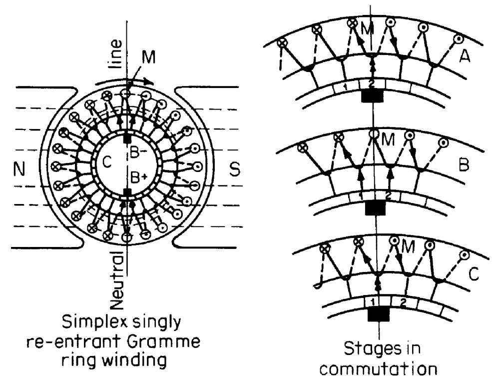


FIGURE 8-8 Principle of commutation.

However, if the coils are connected at a commutator C made up of copper blocks insulated from each other, brushes $B-$ and $B+$ may be used to connect the two halves in parallel with respect to an external circuit and currents will flow in the proper direction in the conductors beneath the poles.

As the armature rotates, the coil M passes from one side of the neutral line to the other and the direction of the current in it is shown at three successive instants at a, b, and c in Fig. 8-9. As the armature moves from a to c and the brush changes contact from segment 2 to segment 1, the current in M is automatically reversed. For a short period, the brush contacts both segments and short circuits the coil. It is important that no voltage be induced in M during that time, or the resulting circulating currents could be damaging. This accounts for the location of the brushes so that M will be at the neutral flux point between the poles.

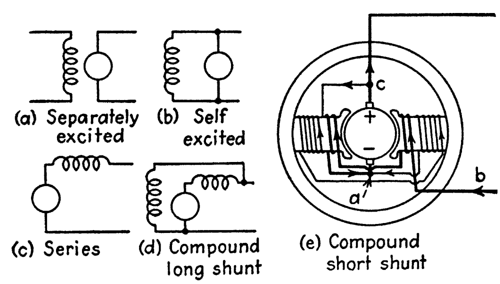


FIGURE 8-9 Methods of excitation.

Field Excitation. Because current-carrying conductors produce flux that links them as described above (in paragraphs on force on current-carrying conductors in a magnetic field), flux from the main poles is obtained by winding conductors around the pole bodies and passing current through them. This current may be supplied in different ways. When a generator supplies its own exciting current, it is "self-excited." When current is supplied from an external source, it is "separately excited." When excited by the load current of the machine, it is "series excited."

8.3 ARMATURE WINDINGS

Terms. The Gramme-ring winding is not used, because half the conductors (those on the inside of the ring) cut no flux and are wasted. Figures 8-8, 8-10, and 8-11 show such windings only because they illustrate types of connections so well.

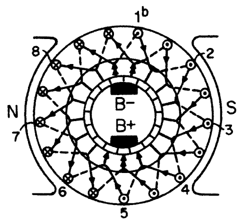


FIGURE 8-10 Singly reentrant duplex winding.

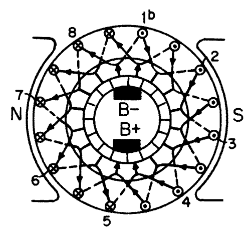


FIGURE 8-11 Doubly reentrant duplex winding.

A *singly reentrant winding* closes on itself only after including all the conductors, as shown in Figs. 8-8 and 8-10.

A *doubly reentrant winding* closes on itself after including half the conductors, as shown in Fig. 8-11.

As shown, a *simplex winding* has only two paths through the armature from each brush (Fig. 8-8). A *duplex winding* has twice as many paths from each brush and is shown in Figs. 8-10 and 8-11. Note that each brush should cover at least two commutator segments with a duplex winding, or one circuit will be disconnected at times from the external circuit. Although it is possible to use multiplex and multiple reentrant windings, they are uncommon in the United States. They are used in Europe in some large machines.

Modern dc machines have the armature coils in radial slots in the rotor. Nonmetallic wedges restrain the coils normally, but some wedgeless rotors use nonmetallic banding around the core, such as glass fibers in polyester resin. This permits shallower slots and helps to reduce commutation sparking. However, the top conductors are near the pole faces and may have high eddy losses. The coil ends outside the slots are held down on coil supports by glass polyester bands for both types.

Multiple, or Lap, Windings. Figure 8-12 shows a *lap-winding coil*. The conductors shown on the left side lie in the top side of the rotor slot. Those on the right side lie in the bottom half of another slot approximately one pole pitch away. At any instant the sides are under adjacent poles, and voltages induced in the two sides are additive. Other coil sides fill the remaining portions of the slots. The coil leads are connected to the commutator segments, and this also connects the coils to form the armature winding. This is shown in Fig. 8-13. The pole faces are slightly shorter than the rotor core.

Almost all medium and large dc machines use simplex lap windings in which the number of parallel paths in the armature winding equals the number of main poles. This permits the current per path to be low enough to allow reasonable-sized conductors in the coils.

Windings. Representations of dc windings are necessarily complicated. Figure 8-14 shows the lap winding corresponding to the Gramme-ring winding of Fig. 8-8. Unfortunately, the nonproductive end portions are emphasized in such diagrams, and the long, useful portions of the coils in the core slots are shown as radial lines. Conductors in the upper layers are shown as full lines, and those

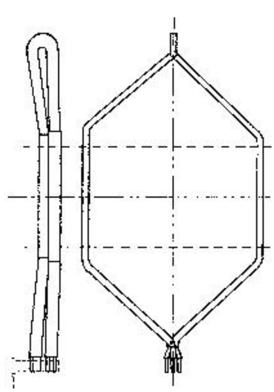


FIGURE 8-12 Coil for one-turn lap winding.

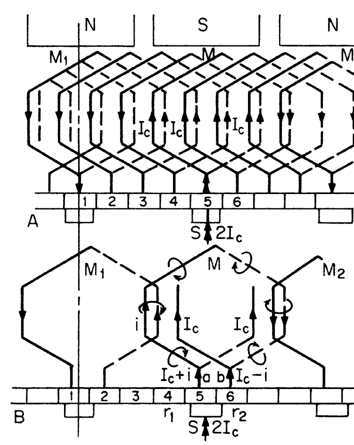


FIGURE 8-13 Multiple, or lap, winding.

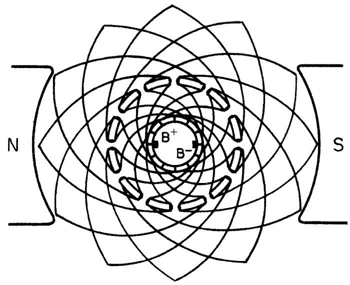


FIGURE 8-14 Simplex lap winding.

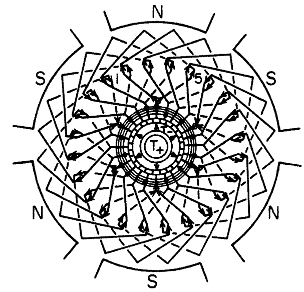


FIGURE 8-15 Simplex singly reentrant full-pitch multiple winding with equalizers.

in the lower layers as dotted lines. The inside end connections are those connected to the commutator bars. For convenience, the brushes are shown inside the commutator.

Note that both windings have the same number of useful conductors but that the Gramme-ring winding requires twice the number of actual conductors and twice the number of commutator bars.

Figure 8-15 shows a 6-pole simplex lap winding. Study of this reveals the six parallel paths between the positive and negative terminals. The three positive brushes are connected outside the machine by a copper ring $T+$ and the negative brushes by $T-$.

The two sides of a lap coil may be full pitch (exactly a pole pitch apart), but most machines use a short pitch (less than a pole pitch apart), with the coil throw one-half slot pitch less than a pole pitch. This is done to improve commutation.

Equalizers. As shown in Fig. 8-15, the parallel paths of the armature circuit lie under different poles, and any differences in flux from the poles cause different voltages to be generated in the various paths. Flux differences can be caused by unequal air gaps, by a different number of turns on the main-pole field coils, or by different reluctances in the iron circuits.

With different voltages in the paths paralleled by the brushes, currents will flow to equalize the voltages. These currents must pass through the brushes and may cause sparking, additional losses, and heating. The variation in pole flux is minimized by careful manufacture but cannot be entirely avoided.

To reduce such currents to a minimum, copper connections are used to short-circuit points on the paralleled paths that are supposed to be at the same voltage. Such points would be exactly two pole pitches apart in a lap winding. Thus in a 6-pole simplex lap winding, each point in the armature circuit will have two other points that should be at its exact potential. For these points to be accessible, the number of commutator bars and the number of slots must be a multiple of the number of poles divided by 2.

These short-circuited rings are called "equalizers." Alternating currents flow through them instead of the brushes. The direction of flow is such that the weak poles are magnetized and the strong poles are weakened. Usually, one coil in about 30% of the slots is equalized. The cross-sectional area of an equalizer is 20% to 40% that of the armature conductor.

Involute necks, or connections, to each commutator bar from conductors two pole pitches apart give 100% equalization but are troublesome because of inertia and creepage insulation problems.

Figure 8-15 shows the equalizing connections behind the commutator connections. Normally they are located at the rear coil extensions, and so they are more accessible and less subject to carbon-brush dust problems.

Two-Circuit, or Wave, Windings. Figure 8-16 shows a wave type of coil. Figure 8-17 gives a 6-pole wave winding. Study reveals that it has only two parallel paths between the positive and negative terminals. Thus, only two sets of brushes are needed. Each brush shorts $p/2$ coils in series. Because points a, b, and c are at the same potential (and, also, points d, e, and f), brushes can be placed at each of these points to allow a commutator one-third as long.

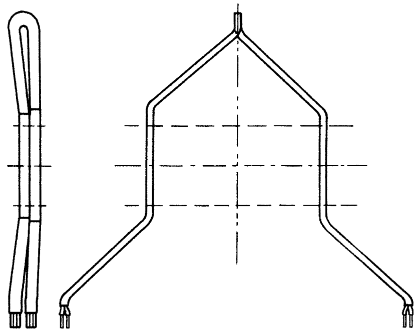


FIGURE 8-16 One-turn wave winding.

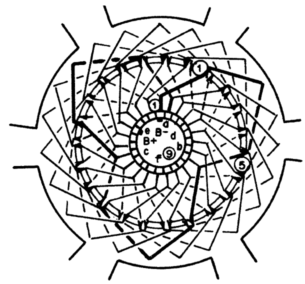


FIGURE 8-17 Two-circuit progressive winding.

The winding must progress or retrogress by one commutator bar each time it passes around the armature for it to be singly reentrant. Thus, the number of bars must equal $(kp/2) \pm 1$, where k is a whole number and p is the number of poles. The winding needs no equalizers because all conductors pass under all poles.

Although most wave windings are 2-circuit, they can be multicircuit, as 4 or 16 circuits on a 4-pole machine or 6, 12, or 24 circuits on a 12-pole machine. Multicircuit wave windings with the same number of circuits as poles can be made by using the same slot and bar combinations as on a lap winding. For example, with an 8-pole machine with 100 slots and 200 commutator bars, the bar throw for a simplex lap winding would be from bar 1 to bar 2 and then from bar 2 to bar 3, etc. For an 8-circuit wave winding, the winding must fail to close by circuits/2 bars, or 4. Thus, the throw would be bar 1 to 50, to bar 99, to bar 148, etc. The throw is (bars $\pm$ circuits/2)$(p/2)$, in this case, $(200 - 4)/4 = 49$. Theoretically such windings require no equalizers, but better results are obtained if they are used.

Since both lap and multiple wave windings can be wound in the same slot and bar combination simultaneously, this is done by making each winding of half-size conductors. This combination resembles a *frog's leg* and is called by that name. It needs no equalizers but requires more insulation space in the slots and is seldom used.

Some wave windings require *dead coils.* For instance, a large 10-pole machine may have a circle of rotor punchings made of five segments to avoid variation in reluctance as the rotor passes under the five pairs of poles. To avoid dissimilar slot arrangements in the segments, the total number of slots must be divisible by the number of segments, or 5 in this case. This requires the number of commutator bars to be also a multiple k of 5. However, the bar throw for a simplex wave winding must be an integer and equal to (bars $\pm$ 1)$(p/2)$. Obviously $(5k \pm 1)/5$ cannot meet this requirement. Consequently one coil, called a *dead coil*, will not be connected into the winding, and its ends will be taped up to insulate it completely. No bar will be provided for it, and thus the bar throw will be an integer. Dead coils should be avoided because they impair commutation.

8.4 ARMATURE REACTIONS

Cross-Magnetizing Effect. Figure 8-18*a* represents the magnetic field produced in the air gap of a 2-pole machine by the mmf of the main exciting coils, and part *b* represents the magnetic field produced by the mmf of the armature winding alone when it carries a load current. If each of the Z armature conductors carries I_c A, then the mmf between *a* and *b* is equal to ZI_c/p At. That between *c* and *d* (across the pole tips) is $\psi ZI_c/p$ At, where ψ = ratio of pole arc to pole pitch. On the assumption that all the reluctance is in the air gap, half the mmf acts at *ce* and half at *fd*, and so the cross-magnetizing effect at each pole tip is

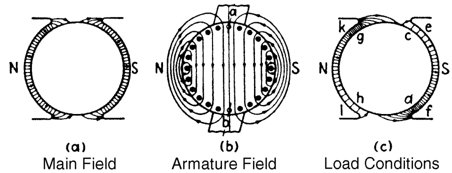


FIGURE 8-18 Flux distribution in (*a*) main field, (*b*) armature field, and (*c*) load conditions.

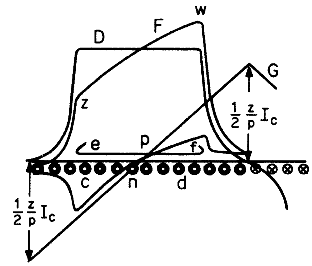


FIGURE 8-19 Flux distribution in a large machine with *p* poles.

$$\frac{\psi Z I_c}{2p} \quad \text{ampere-turns} \tag{8-1}$$

for any number of poles.

Field Distortion. Figure 8-18*c* shows the resultant magnetic field when both armature and main exciting mmfs exist together; the flux density is increased at pole tips *d* and *g* and is decreased at tips *c* and *h*.

Flux Reduction Due to Cross-Magnetization. Figure 8-19 shows part of a large machine with *p* poles. Curve *D* shows the flux distribution in the air gap due to the main exciting mmf acting alone, with flux density plotted vertically. Curve *G* shows the distribution of the armature mmf, and curve *F* shows the resultant flux distribution with both acting. Since the armature teeth are saturated at normal flux densities, the increase in density at *f* is less than the decrease at *e*, so that the total flux per pole is diminished by the cross-magnetizing effect of the armature.

Demagnetizing Effect of Brush Shift. Figure 8-20 shows the magnetic field produced by the armature mmf with the brushes shifted through an angle θ to improve commutation. The armature field is no longer at right angles to the main field but may be considered the resultant of two components, one in the direction *OY*, called the "cross-magnetizing component," and the other in the direction *OX*, which is called the "demagnetizing component" because it directly opposes the main field. Figure 8-21 gives the armature divided to show the two components, and it is seen that the demagnetizing ampere-turns per pair of poles are

$$\frac{Z I_c}{p} \times \frac{2\theta}{180} \quad \text{ampere-turns} \tag{8-2}$$

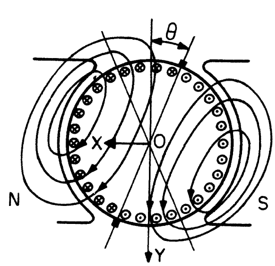


FIGURE 8-20 Demagnetizing effect.

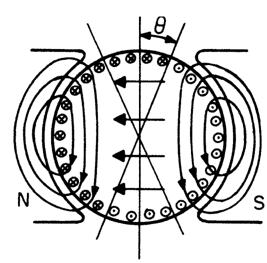


FIGURE 8-21 Cross-magnetizing effect.

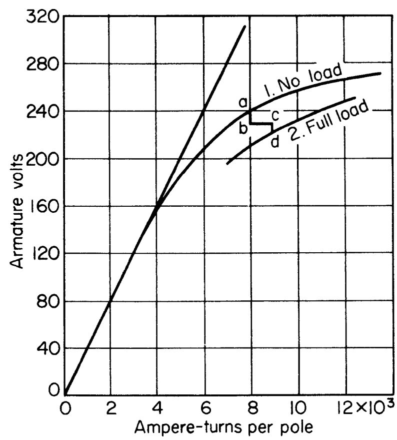


FIGURE 8-22 Saturation curves—dc generator.

where $2\theta/180$ is about 0.2 for small noncommutating pole machines where brush shift is used. The demagnetizing ampere-turns per pole would be

$$0.1ZI_c/p \quad \text{ampere-turns} \tag{8-3}$$

No-Load and Full-Load Saturation Curves. Curve 1 of Fig. 8-22 is the no-load saturation curve of a dc generator. When full-load current is applied, there is a decrease in useful flux, and therefore a drop in voltage *ab* due to the armature cross-magnetizing effect (see paragraph on flux reduction, above). A further voltage drop from brush shift is counterbalanced by an increase in excitation $bc = 0.1\ ZIc/p$; also a portion *cd* of the generated emf is required in overcoming the voltage drop from the current in the internal resistance of the machine. The no-load voltage of 240 V requires 8000 At. At full load at that excitation the terminal voltage drops to 220 V. To have both no-load and full-load voltages equal to 240 V, a series field of 10,700 − 8000 = 2700 At would be required.

8.5 COMMUTATION

Commutation Defined. The voltages generated in all conductors under a north pole of a dc generator are in the same direction, and those generated in the conductors under a south pole are all in the opposite direction (Fig. 8-23). Currents will flow in the same direction as induced voltages in generators and in the opposite direction in motors. Thus, as a conductor of the armature passes under a brush, its current must reverse from a given value in one direction to the same value in the opposite direction. This is called "commutation."

Conductor Current Reversal. If commutation is "perfect," the change of the current in a coil will be linear, as shown by the solid line in Fig. 8-24. Unfortunately, the conductors lie in steel slots, and self-and mutual inductances in Fig. 8-25 cause voltages in the coils short-circuited by the brushes. These result in circulating currents that tend to prevent the initial current change, delaying the reversal. In extreme cases, the delay may be as severe as indicated by the dotted line of Fig. 8-24. Because the current must be reversed by the time the coil leaves the brush (when there is no longer any path for circulating currents), the current remaining to be reversed at *F* must discharge its energy in an

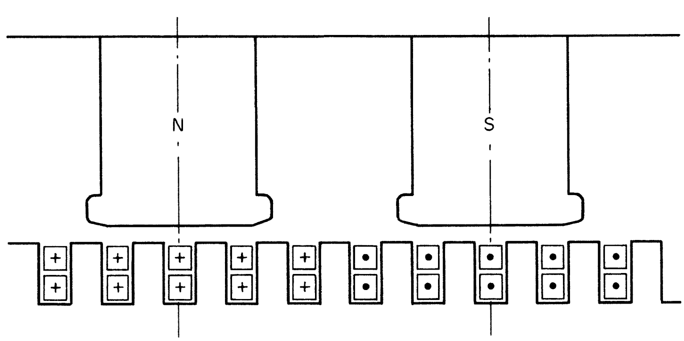


FIGURE 8-23 Conductor currents.

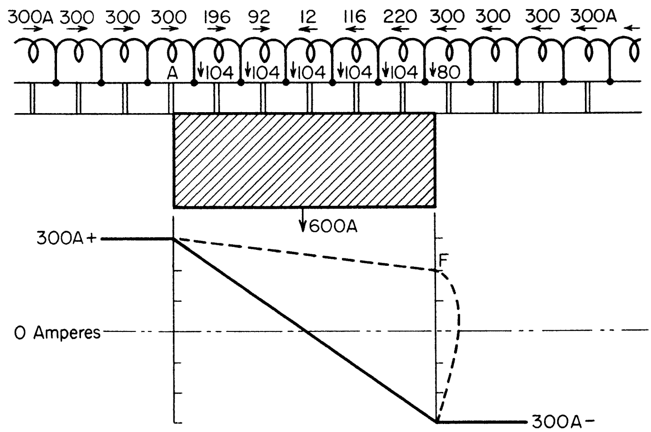


FIGURE 8-24 Commutation.

electric arc from the commutator bar to the heel of the brush. This is commutation sparking. It can burn the edges of the commutator bars and the brushes. However, most large and heavy-duty dc machines have some nondamaging sparking, and "sparkless" commutation is not required by accepted standards. However, commutation must not require undue maintenance.

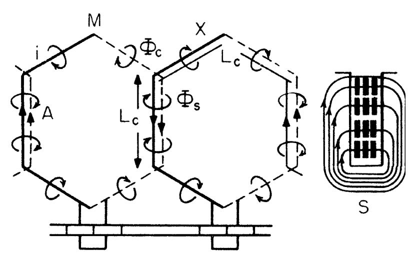


FIGURE 8-25 Magnetic field surrounding short-circuited coils.

The undesired voltages causing the circulating currents result from interpolar fluxes from armature reaction, leakage fluxes of the current-carrying armature conductors, and, in some cases, main-pole-tip spray flux. Beneficial factors reducing the circulating currents include the resistance of the short-circuited coil, the resistance of the commutator risers, and that of the brush body to transverse currents. However, the most important factor is the voltage drop at the sliding contact between the brush face and the copper commutator surface.

Commutator Brushes. Most dc machines use electrographitic brushes with about 60 A/in^2 current density at full load. These have an essentially constant contact voltage drop at the commutator surface of about 1 V for loads above one-third. This effective resistance to circulating currents is important to good operation of dc machines.

The cross-resistance of the brush body to circulating currents can be increased by splitting the brush into two wafers and making the crosscurrents cross the air gap between the two pieces. This has increased the good commutation range on some machines by 7%. The use of double brush holders, which have metal dividers between two brushes in the holder, is even more effective and has increased the good commutation range as much as 15% over single solid brushes.

Unless special brushes are used, machines should be operated for not more than a few hours at a time at brush densities below 30 A/in^2. If this is done, the commutator surface develops a hard glaze which makes the brushes chatter. This results in frayed shunts, chipped and broken brushes, and excessive brush-finger wear.

Reactance Voltage of Commutation. The sum of the voltages induced in the armature coil while it is short-circuited by the brushes while undergoing commutation is called the *reactance voltage of commutation*. One of the most important of the fluxes causing this voltage is the *slot-leakage flux* shown in

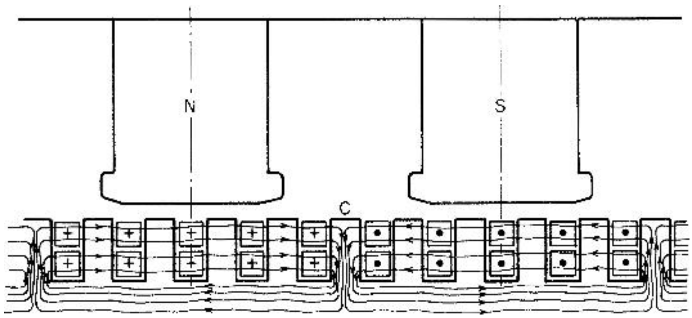


FIGURE 8-26 Slot-leakage flux.

Fig. 8-26. This is the resultant flux leakage from current in the individual slot conductors, as shown in Fig. 8-25. Because the radial fluxes in the rotor teeth from adjacent slot conductors essentially cancel except at point *C* (the point of current reversal), the resultant flux is as shown in Fig. 8-26. As the conductors commutate and pass through *C*, they cut the flux shown there and this generates the *reactance voltage of commutation.* Actually, part of this voltage is also due to leakage-flux changes at the coil ends, to armature reaction flux, etc., but, for simplicity, only the important slot leakage flux is shown.

Commutating Poles. The beneficial factors that limit the circulating currents in coils being commutated are not adequate to prevent serious delays in current reversal. Other means must be taken to prevent sparking.

If the flux at *C* (Fig. 8-26) could be nullified by an equal flux in the opposite direction, the circulating currents due to the slot leakage flux would be prevented.

The location of *C* is fixed by the location of the brushes. If the brushes were shifted toward the south main pole, a position could be found where the main flux upward into the south pole would cancel the downward flux due to slot leakage at *C*.

This method was used in the early history of dc machines. Unfortunately, the slot-leakage flux at *C* is proportional to conductor load current, whereas the flux into the south pole is not. Thus, a new brush position is needed for every change in load current.

A better solution is to provide stationary poles midway between the main poles, as shown in Fig. 8-27. Windings on these *commutating poles* carry the load current. Thus, the flux into the pole at *C* is proportional to the rotor conductor currents and, theoretically, can cancel the voltages induced in the coils being commutated by the slot leakage flux. In the case of the dc motor, the current

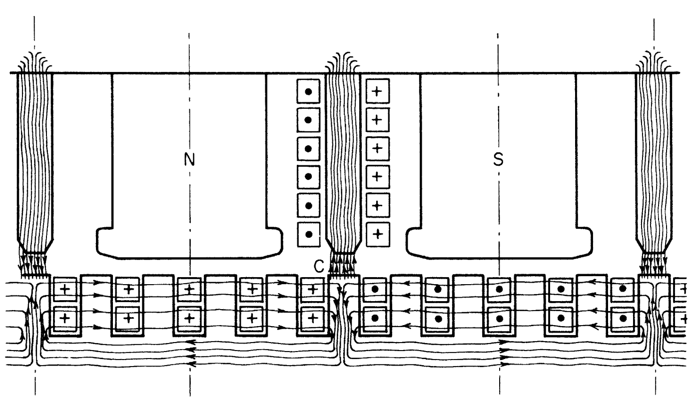


FIGURE 8-27 Slot-leakage flux and commutating-pole flux.

reverses in both the armature and the commutating field, and proper canceling is maintained.

Note that the strength of the commutating-pole winding must be greater than the armature-winding ampere-turns per pole by the amount required to carry the needed flux across the commutating-pole air gap.

Almost all modern dc machines use commutating poles, although some small machines have only half as many as main poles.

The commutating-pole tip is usually shaped with tapered sides to approximate the shape of the reactance voltage of commutation form (see Figs. 8-27 and 8-28).

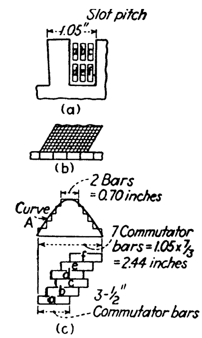


FIGURE 8-28 Commutating zone.

Reactance Voltage of Commutation Formula. To determine the useful flux needed across the commutating-pole air gap, it is useful to calculate the reactance voltage of commutation (the total of the voltages induced in the armature coil as it undergoes commutation). The approximate value of this voltage may be calculated by the use of the following formula:

$$E_c = \frac{\text{poles}}{\text{paths}}(l_c ZT)(\text{r/min})(10^{-10})\left[(K_i L_r) + K_2(PP)(4.5 + 0.2t_s) + \frac{L_r}{b_s}(3d_s + 2SP)\right] \text{ volts} \tag{8-4}$$

where
- I_c = current per armature conductor, A
- Z = total no. of armature conductors
- T = no. of turns/coil between commutator bars
- L_r = gross armature-core length, in
- K_1 = 18.5 for noncommutating-pole machines
  = 0 for machines with commutating-pole length = L_r
- K_2 = 1.0 for machines using nonmagnetic bands
  = 1.7 for machines using magnetic bands
- PP = pole pitch, in
- t_s = coil throw, slots
- b_s = width of slot, in
- d_s = depth of slot, in
- SP = slot pitch, in

This formula is based on the work by Lamme. (See Theory of Commutation by B. G. Lamme, *Trans. AIEE*, Oct. 1911, vol. 30.)

The Commutating Zone. This is defined as that space on the armature periphery through which a given slot moves while all the conductors lying in the slot commutate. In chorded windings, it is extended to include the coil edges in the chorded slots. The commutating zone thus depends on the number of commutating bars covered per brush.

The zone may be calculated by the following formula:

$$CZ = \frac{SP[(B/S) + (B/S \times Ch) + (B/Br) - Cir/p)]}{B/S} \tag{8-5}$$

where CZ is the commutating zone in inches, SP the rotor slot pitch in inches, B/S the number of commutator bars per slot, Ch the slot chording as a fraction of the slot pitch, B/Br the number of commutator bars spanned per brush, Cir the number of paralleled circuits in the armature, and p the number of main poles.

Consider an 8-pole simplex lap winding with three bars per slot, chording of ½ slot, 3½ bars per brush, and slot pitch of 1.05 in:

$$CZ = \frac{1.05 \times (3 + 1\tfrac{1}{2} + 3\tfrac{1}{2} - 8/8}{3} = 2.44 \text{ in}$$

In this machine, all the conductors in a slot are commutated while the armature periphery moves 2.44 in.

This can be seen graphically in Fig. 8-28, where (*a*) shows a slot with six conductors, (*b*) shows a brush covering 3½ bars, and (*c*) shows the graphical solution. In (*c*) the rectangle *a* represents as abscissa the space of 3½ commutator bars if they were at the armature surface. This is the length to commutate coil *a*. The ordinate represents to a convenient scale the commutation voltage induced in this conductor while it is being commutated. Rectangles *b* and *c* are the same for coils *b* and *c*. Since *b* commutates 1 bar later than *a*, it is shown one bar space to the right of *a*, etc. In a similar manner *d*, *e*, and *f* are shown. Normally *d* would be expected to start commutation at the same time as *a*, but, because of chording, it starts later, in this case 1½ bars later. Thus, the commutating zone starts with the beginning of rectangle *a* and is completed at the end of rectangle *f*. On adding the spaces of the parts, this is 3½ bars for *f*, 2 bars for the steps of *e* and *d*, and 1½ bars for chording, or a total of 7 bars at the rotor surface, which is 1.05 7⁄3, or 2.44 in.

The summation of the individual rectangles as smoothed off by curve *A* of (*c*) is a rough representation of the reactance voltages induced in the coils during commutation.

Single Clearance. The centerline of the commutating zone and curve *A* of Fig. 8-28 lie midway between the adjacent main-pole tips if the brushes are not shifted off neutral. The arc on the rotor surface between the tips of adjacent main poles is called the *neutral zone*. If the commutating zone is centered in this arc, the spaces left at each end are called the *single clearance*. Thus, the single clearance is

$$SC = (\text{neutral zone} - \text{commutating zone})/2 \tag{8-6}$$

The single clearance is an indication of the probability that spray flux from the main-pole tips might flow into the commutating zone. Such flux would not vary with load and would distort the form of the useful flux from the commutating pole. The commutating-pole useful flux form should closely resemble that of curve *A* in Fig. 8-28.

Noncompensated dc machines usually have main-pole tips with short radial dimensions and have limited spray flux into the neutral zone. The minimum single clearance for these should be not less than 0.6 in and not less than 0.9 in with commutation voltages above 3 or 4 V.

Compensated-machine main poles usually have tips 2 to 3 in deep to accommodate the compensating slots and are more likely to spray flux into the commutating zone. These require single-clearance minimums of 1.2 to 1.4 in.

If there is any question about tip flux reaching the commutating zone, flux plots should be made.

Commutating-Pole Excitation. Figures 8-18*b* and 8-19 show that flux should normally be expected in the commutation area. It is caused by the armature-winding ampere-turns per pole. It could be reduced to zero if the commutating pole had ampere-turns equal and opposite to those of the armature winding. This is $ZI_c/2p$ At/pole.

However, it is necessary that the commutating winding also produces useful flux across the commutating-pole gap to counteract the reactance voltage of commutation, as shown in Fig. 8-27. For this reason, the strength of the commutating field is usually 20% to 30% greater than the armature ampere-turns per pole. This difference is called the *excess ampere-turns*. These must be added to the circled dotted-line bar diagram of Fig. 8-29. The actual flux across the gap is set accurately during the factory test by adjusting the number of sheet-steel shims behind the commutating poles to set the reluctance of the gap for the exact flux needed.

Calculation of Commutating-Pole Air Gaps. With fixed excess ampere-turns on the commutating-pole winding and a certain commutation voltage at rated current and speed, only one particular commutating-pole air gap will result in the most favorable compensation of the commutation voltage. The shape of the pole tip will determine the form of the flux density under it, but the length of the air gap will determine the magnitude of the density.

To counteract the reactance voltage of commutation E_c, the approximate maximum flux density needed in the commutating-pole air gap is

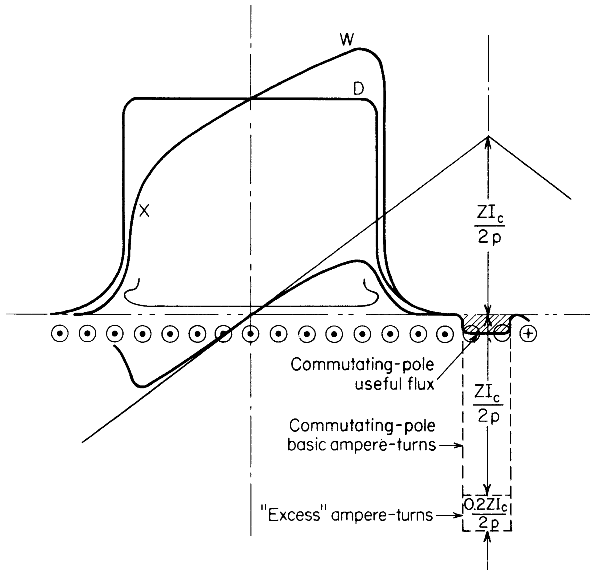


FIGURE 8-29 Commutating-pole ampere-turns.

$$B_m = \frac{E_c \times 23 \times 10^8}{Z/\text{bar} \times DL_c \times \text{r/min}} \tag{8-7}$$

where E_c is the full-load reactance voltage of commutation at speed r/min, Z is the total number of armature conductors, bars is the total number of commutator bars, D is the armature diameter in inches, L_c is the axial length of the commutating poles in inches, and r/min is the revolutions per minute for which E_c was calculated.

The approximate length of the needed commutating-pole single air gap may be calculated by the following formula:

$$\text{Gap} = \frac{3.19 \times \text{excess ampere-turns}}{B_m} \tag{8-8}$$

When the machine is on factory test, the excess ampere-turns can be adjusted to obtain the best commutation possible by placing another dc generator or a battery across the commutating winding to add to the load current flowing in it or to lower the excess by shunting out some of the load current. This is known as a "boost or buck" test. Afterward the commutating-pole air gap is changed to produce the "best" gap flux density with the actual excess ampere-turns. The new gap will be

$$\text{Gap}_2 = \frac{\text{excess At}_1}{\text{excess At}_2} \times \text{gap}_1 \tag{8-9}$$

Dimensions of Commutating Poles. If the useful flux across a commutating-pole air gap is not proportional to the machine load current, the compensation of the reactance voltage of commutation will not be correct for all loads and sparking may damage the brushes and commutator. Thus, the commutating pole must not saturate at the highest load currents to be accommodated. The base of the pole must carry not only the useful air-gap flux but also leakage fluxes from the commutating and main field coils which are near. These leakage fluxes are relatively large and must be determined with care by flux plotting if the danger of commutating-pole saturation exists.

The amount of leakage flux through the base of the pole depends on the length of the leakage paths, the number of coil ampere-turns, and the location of the commutating field. The leakage paths should be made as long as feasible, the coil ampere-turns as few as reasonable, and the commutating coil located as close to the pole tip as possible. Also, all sections of the commutating pole should be large enough to accommodate their flux.

For a normal compensated machine, the leakage flux will be about 75% of the commutating-pole useful flux, or about 140% of the useful flux in a noncompensated machine.

The approximate useful flux can be calculated by using the maximum commutating-pole air-gap flux density from Eq. (8-7). The average flux density of the commutating zone will be approximately

$$B_a = 0.83B_m \tag{8-10}$$

The flux density at overload in the base of the pole is

$$B_{cp} = \frac{K_3 \times K_4 \times B_a \times CZ}{L_c \times W_c} \tag{8-11}$$

where K_3 is 1.75 for compensated machines and 2.40 for noncompensated machines, K_4 is the ratio of overload current to rated current, B_a is the average flux density in the commutating zone, CZ is the width of the commutating zone, L_c is the axial length of the commutating pole, and W_c is the circumferential width of the pole at its base. B_{cp} should not exceed 80,000 to 90,000 lines/in^2 for good commutation.

Compensating Windings. Although the commutating pole is a good solution for commutation, it does not prevent distortion of the main-pole flux by armature reaction. The flux set up across the main-pole face by the armature mmf is shown in Fig. 8-30*a*. If the pole face is provided with another winding, as shown in Fig. 8-30*b*, and connected in series with the load, it can set up an mmf equal and opposite to that of the armature. This would tend to prevent distortion of the air-gap field by armature reaction. Such windings are called *compensating windings* and are usually provided on medium-sized and large dc machines to obtain the best possible characteristics. They are also often needed to make machines less susceptible to flashovers.

The use of compensating windings reduces the number of turns required on the commutating-pole fields, and this materially reduces the leakage fluxes of the field and, in turn, the pole saturations at high currents. The ampere-turns on the commutating field are reduced by about 50% with the use of a compensating field. This new winding may be considered to be some of the turns taken off the commutating-pole winding and relocated in slots in the main-pole faces.

The number and location of the compensating slots must be carefully chosen to match, as closely as possible, the rotor ampere-turns per inch. However, the slot spacing must not correspond closely to that of the rotor. This would cause a major change in reluctance to the main-pole useful flux every time the rotor moved from a position where the rotor and stator slots all coincided to where the rotor slots coincided with the stator teeth. This would occur once for every slot-pitch movement. The resulting rapid changes in useful flux would cause ripples in the output voltage and also serious magnetic noise. If too few slots are used, local flux distortions occur and the compensating winding loses some of its effectiveness (see Fig. 8-32).

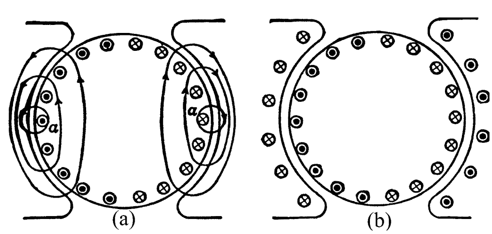


FIGURE 8-30 Armature field without (*a*) and with (*b*) compensating windings.

Compensation of armature reaction effectively reduces the armature circuit inductance. This makes the machine less susceptible to the bad effects of $L(di/dt)$ voltages caused by very fast load current changes.

During manufacture, it is possible to locate the compensating winding nonsymmetrically about the centerline of the main pole. This causes a direct-axis flux, which will give a series field effect (Fig. 8-31). For generator cumulative compounding, the slots must be shifted in the direction of the machine rotation. This shift gives a motor differential compounding. The effect cannot be adjusted after manufacture. It seldom exceeds ½ in, and this does not materially reduce the effectiveness of the compensation.

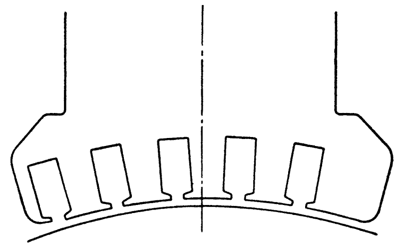


FIGURE 8-31 Offset compensating winding.

Volts per Bar. The mica thickness between the commutator segments depends on the machine design and varies from 0.020 in on small machines to 0.050 in on large units. Although several hundred volts would normally be required to jump these distances, the presence of ionized air from sparking and the presence of conducting carbon dust make it necessary that the voltage between segments be held to low values. If a low-resistance arc does jump between segments, it raises the voltages across the remaining bars. It also tends to ionize some air to form conducting paths across the rest of the bars. If this progresses until all the segments between brush arms of opposite polarity are bridged, then a *flashover* occurs and severe damage may result to the commutator, brushes, and brush holders.

Because the highest voltage between bars is the "trigger" that starts the flash, this is an important limit. The "average" volts per bar has little significance. Figure 8-29 shows that the maximum volts per bar depends on the field form. For the noncompensated machine shown, the maximum volts between segments exists at w. The segments connected to conductors at x have much less voltage between them, and those beyond the edge of the pole have almost none.

The relation between maximum volts per bar and the average depends on the armature ampere-turns per pole and the saturation curve of the gap and teeth at the pole tips. On neglecting the small voltage drop in the series and commutating windings, the voltage between brush arms is the machine voltage V, and the number of bars between arms is B/p. Thus

$$\text{Average volts/bar} = \frac{V \times p}{B} \tag{8-12}$$

where B is the total number of commutator bars and p the number of main poles.

Even if no distortion exists, only the conductors under the pole faces generate voltage, and so the corrected average volts per bar should be

$$\frac{V \times p}{B \times \psi} \quad \text{volts}$$

where ψ is the ratio of pole arc to pole pitch, about 0.65. This is represented by D in Fig. 8-29. However, the maximum volts per bar at w is greater than this, as the height w is greater than D, or

$$\text{Maximum volts per bar} = \frac{V \times p}{B} \times \frac{w}{D \times \psi} \tag{8-13}$$

In practice, the value of w/D for a noncompensated machine at full-strength main field varies from about 1.7 to 1.9. However, any reduction in saturation causes the effects of the armature ampereturns (which cause the distortion) to be magnified. The designer must check the actual value of w/D, since it may be as high as 4.5 for a dc motor at a weak main field strength (high speed). This is evident in Fig. 8-32. The distorting effect for the high-speed (low-average-flux) condition ϕ_{02} raises the maximum flux to ϕ_{w2}, which is over 3 times the change for the saturated

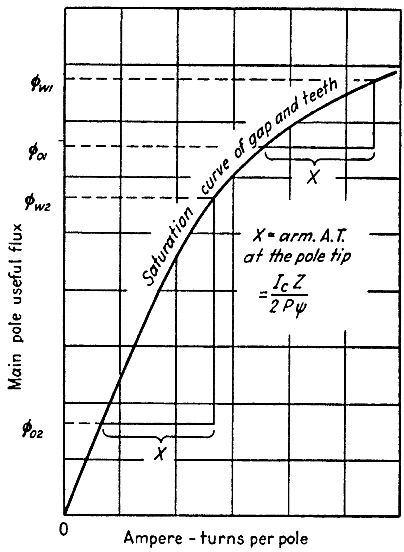


FIGURE 8-32 Effect of flux-distortion armature ampere-turns at normal and low saturation.

(low-speed) condition ϕ_{01} to ϕ_{w1} with the same distorting ampere-turns X.

The use of a compensating winding tends to eliminate the flux distortion, and for saturated conditions the flux curve coincides well with the no-load curve D of Fig. 8-29. However, under low saturation conditions the stationary compensating windings permit localized flux distortions. These are shown in Fig. 8-33. Similar distortions occur at low main flux densities on dc generators, but the output voltage V is reduced in the same proportion as the main flux, and the maximum voltage between bars is not affected seriously.

At full field on well-compensated motors or generators *w/D* is about 1.4 to 1.5. Direct-current motors at weak field may have ratios of 2.0 or more. On any questionable machine the designer should check this value carefully.

Approximate safe limits of maximum volts per bar are 40 V for motors and 30 V for generators on machines having 0.040-in-thick mica between segments.

Brush Potential Curves. When a dc machine develops some commutation sparking, the user may suspect that the commutating-pole air gap is not set correctly. "Brush potential curves" are often taken to prove or disprove such suspicions.

These are taken by measuring the voltage drops between the brush and commutator surface at four points while the machine is operating at constant speed and load current (see Fig. 8-34). The voltages at 1, 2, 3, and 4 are taken by touching the pointed lead of a wooden pencil to the commutator surface. The circuit is completed with leads and a low-reading voltmeter is shown.

The voltages are then plotted. A curve such as A of Fig. 8-34 may indicate undercompensation due to a too large commutating-pole gap. Curve C may indicate overcompensation with too much flux density in the commutating-pole air gap. Curve B is typical of good compensation.

Justification for such conclusions is based on the theory that best commutation (coil current reversal) will be linear while the coil passes under the brush. This is possible only if there are no circulating currents. Undercompensation should cause circulating currents that would crowd the current to the leaving edge of the brush and cause a high voltage at point 4. Overcompensation would reverse the current too soon and would actually reverse the voltage drop at point 4.

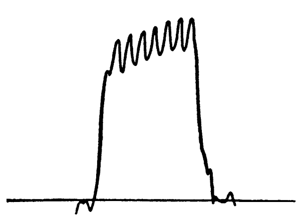


FIGURE 8-33 Main-pole flux distortion on a compensated motor at full load and $2\frac{1}{2}$ times base speed.

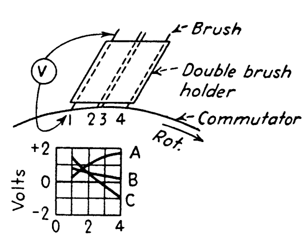


FIGURE 8-34 Brush potential curves.

Even to an expert, this test is only an indicator that more definitive tests, such as a buck-boost test, are needed [see Eq. (8-8)]. Many other factors, including brush riding, commutator surface conditions, and sparking, influence the readings. Where machine changes may be required, the manufacturer should be consulted.

8.6 ARMATURE DESIGN

EMF Equation. If 10^8 lines (Mx) of flux are cut by one conductor in 1 s, 1 V is induced in it. Therefore, the induced voltage of a dc machine is

$$E = \phi_t \times \frac{Z}{C} \times \frac{\text{r/min}}{60} \times 10^{-8} \tag{8-14}$$

where ϕ_t is the total flux in maxwells across the main air gaps and Z/C is the number of conductors in series per circuit (C).

Output Equation. Equation (8-14) is converted to watts output if both sides are multiplied by the load current I_L, $I_c \times C$. The formula can then be rearranged as

$$D^2L = \frac{\text{watts} \times 6.08 \times 10^8}{\text{r/min} \times B_g \times \psi \times q} \tag{8-15}$$

where D is the armature diameter and L is the armature gross core length, B_g is the main-pole air-gap density in maxwells (lines), ψ is the ratio of pole arc to pole pitch, q is $ZI_c/\pi D$ (a useful loading factor), and ϕ_t is the total air-gap flux equal to

$$B_g \psi \pi D L \tag{8-16}$$

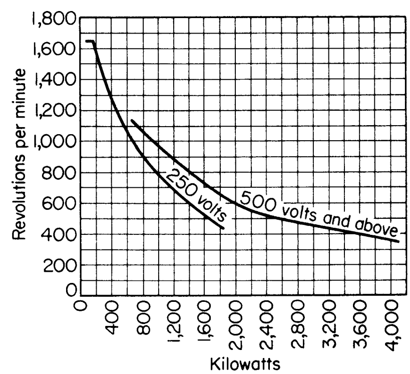


FIGURE 8-35 Standard speeds of dc generators.

Rotor Speeds. Standards list dc generator speeds as high as are reasonable to reduce their size and cost. This relation is seen from Eq. (8-15). The speeds may be limited by commutation, maximum volts per bar, or the peripheral speeds of the rotor or commutator. Generator commutators seldom exceed 5000 ft/min, although motor commutators may exceed 7500 ft/min at high speeds. Generator rotors seldom exceed 9500 ft/min. Figure 8-35 shows typical standard speeds. If the prime mover requires lower speeds than these, generators can be designed for them but larger machines result.

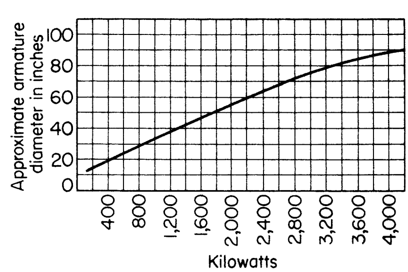


FIGURE 8-36 Approximate rotor diameters for standard speeds of dc generator.

Rotor Diameters. Difficult commutating generators benefit from the use of large rotor diameters, but diameters are limited by the same factors as rotor speeds listed above. The resultant armature length should be not less than 60% of the pole pitch, because such a small portion of the armature coil would be used to generate voltage. Typical generator diameters are shown in Fig. 8-36.

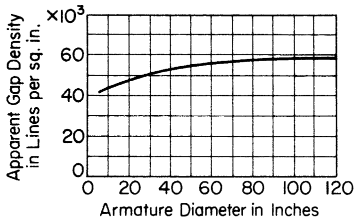


FIGURE 8-37 Curve of apparent gap density versus armature diameter.

Direct-current motor speeds must suit the application, and often the rotor diameter is selected to meet the inertia requirements of the application. Core lengths may be as long as the diameter. Such motors are usually force-ventilated.

Number of Poles and Other Rotor Design Factors. The rotor diameter usually fixes the number of main poles. Typical pole pitches range from 17.5 to 20.5 in on medium and large machines. When a choice is possible, high-voltage generators use fewer poles to allow more voltage space on the commutator between the brush arms. However, high-current generators need many poles to permit more current-carrying brush arms and shorter commutators. Commutators for 1000 to 1250 A/(brush arm)(polarity) are costly, and lower values should be used where existing dies will permit.

The main-pole air-gap flux density Bg is limited by the density at the bottom of the rotor teeth. The reduced taper in the teeth of large rotors permits the higher gap densities, as shown in Fig. 8-37.

Ampere conductors per inch of rotor circumference (q) is limited by rotor heating, commutation, and, at times, saturation of commutating poles. Approximate acceptable values of q are shown in Fig. 8-38.

The commutator diameter is usually about 55% to 85% of the rotor diameter, depending on the sizes available to the designer, the peripheral speed, and the resulting single clearances. Heating may also limit the choice.

Brushes and brush holders are chosen from designs available to limit the brush current density to 60 to 70 A/in^2 at full load, to obtain the needed single clearance, and to obtain acceptable commutator heating.

Selection of an Approximate Design. Consider a generator rated 2500 kW, 700 V, 3571 A, and 514 r/min. From Figs. 8-38 and 8-39

Approx. dia.	$D = 62$ in
Available dia.	$D = 56$ in
No. of poles	10
Pole pitch	17.59 in
Pole arc	12.0 in
(Arc/pitch)	0.687
Neutral zone	6.04 in
B_g gap density at 721 V	58,500 lines/in^2
Approx. q (Fig. 8-38)	1480 A cond./in
D^2L [Eq. (8-15)]	50,200 in^3
L (gross core)	16 in
No. of $\frac{3}{8}$-in vents in core	5
Net core length	14.125 in
ϕ [Eq. (8-14)]	1.12×10^8
Approx. total cond. Z	752 (use 750)
Actual q	1520 A cond./in
No. of commutating bars (1-turn lap)	375
No. of slots	125
Slot pitch	1.407 in
Slot throw	$12\frac{1}{2}$ (use 12)
Chording	$\frac{1}{2}$-slot pitches

Examination of the data indicates that the design appears feasible, and so we may continue.

Commutating dia.	39 in
Brush size (35°)	2(0.500 × 1.75) in
CZ (commutating zone) [Eq. (8-5)]	3.53 in
Brushes/arm	7*b*/*a*
Commutating speed	5250 ft/min
Commutating bar pitch	0.327
Brush arc	1.315 in
Bars/arc	4.02 bars
SC [Eq. (8-6)]	1.26 in
Brush density	58.3 A/in^2
Brush I^2R [Eq. (8-29)]	7142 W
Length of commutating face	7(1.75 + 0.063) + 1 = 14.56 in
Brush friction loss [Eq. (8-33)]	6760 W
Watts/in^2 of commutating surface	7.8 W/in^2

Examination of these data also indicates that the proposed design is reasonable.

Armature Slots and Coils. The depth of an armature slot is limited by several factors, including the tooth density, eddy losses in the armature conductors, available core depths, and commutation. For reasonable frequencies (up to 50 Hz on medium and large dc machines), slots about 2 in deep can ordinarily be used.

Acceptable slot pitches range from 0.75 to 1.5 in. Small machines have shallower slots and a lower range of slot pitches. For medium and large machines, a reasonable tooth density usually results if the ratio of slot width to slot pitch is about 0.4.

Eddy losses in the conductors can be large compared with their load I^2R losses. Sometimes these must be reduced by making each armature conductor from several strands of insulated copper wire. The number of strands and their size depend on the frequency and the total depth of the conductor. An approximate formula for reasonable eddy losses is

$$\text{No. of strands} = (0.168)\,(f^{0.83})(d_c^{0.4}) \qquad (8\text{-}17)$$

where f is the frequency in hertz, (r/min × poles)/120, and d_c is the total depth of a conductor.

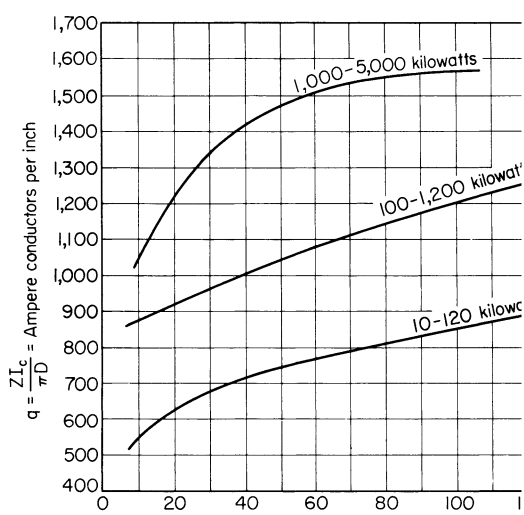


FIGURE 8-38 Ampere conductors per inch of armature circumference.

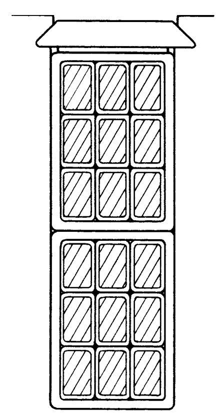


FIGURE 8-39 Armature slot cross section.

The insulation space required depends on the type used. Typical conductor strands have about 0.018 in of glass strands and varnish total. Mica wrappers, binding tapes, and varnish and slot finish allowance (0.010 in) total about 0.085 in on the coil width. If the space for the wedge and its retainer is included, the two coils depthwise total about 0.315 in (see Fig. 8-39).

Approximate Slot Design

Width [see text preceding and following Eq. (8-17)] 0.4 × 1.407	0.563 in
Depth	2.0 in
Approx. total cond. depth	0.875 in
Frequency	42.8 Hz
No. of strands/conductor [Eq. (8-17)]	3

	Slot width, in	Depth, in
Approx.		
Size	0.563 in	2.000 in
Insulation	0.139 (0.085 + 0.054)	0.423 (0.315 + 0.108)
Bare copper	0.424 in	1.577 in
Strand size	0.141 in	0.263 in
Use	3 (0.144	0.289) in strands/conductor
Use available slot	0.570 in	2.250 in

8.7 COMPENSATING AND COMMUTATING FIELDS

Compensating Winding Data. The compensating winding should closely match the armature ampere-turns per inch, should avoid causing magnetic noise, and should result in an acceptable maximum volts per bar. Machines for 40°C temperature rise will have compensating bar densities of about 2500 to 3000 A/in^2. The pole tip section will limit the maximum depth of the compensating bar. Localized areas of high flux density must be avoided where flux must funnel between the pole "shoe" surface and the bottom of the compensating slot.

For single compensating bar-per-slot designs, the typical width required for insulation, varnish, and stacking factor is about 0.140 in. With the wedge space included, the insulation-depth requirement is about 0.400 in.

Compensating Winding Calculations

q (armature)	1520 A cond./in
Pole arc of 12.1 in covers	18,400 A cond.
Approx. compensating At	9200 At
Load current	3571 A
Approx. turns/pole 2.68	Use 2.5 turns/pole
Consider 5 slots/pole	1 bar/slot = 2.5 turns/pole
Size of compensating bar	0.688 × 2.0 in
Bar density	2590 A/in^2
Compensating slot width 0.828	Use 0.830 in
Compensating slot depth 2.400	Use 2.400 in
Compensating slot pitch (layout)	2.25 in
Rotor slot pitch	1.407 in
No magnetic noise	Improbable
Maximum volts/bar	See last two paragraphs in Sec. 8.8.

Commutating Winding Calculations. The total of the commutating and compensating ampere-turns per pole should be about 120% to 130% of those on the rotor.

Armature At/pole = $ZI_c/2p$ = (750 × 357)(2 × 10)	13,400 At/pole
Equiv. armature-turns/pole on line ampere basis	3.75 turns/pole
Approx. commutating + compensating At/pole (1.2 × 13,400)	16,100 At/pole
Commutatingc + ompensating turns/pole 16,100/3571	4.5 turns/pole
Less compensating turns/pole	−2.5 turns/pole
Requires commutating winding of	2.0 turns/pole
Excess At/pole, 16,100–13,400	2700 At/pole

Well-ventilated commutating coils may have densities of 2000 to 2500 A/in^2 (see Fig. 8-48).

Commutation Calculations

E_c = reactance voltage of commutation	5.42 V [see Eq. 8-4)]
Commutating-pole gap density B_m	13,550 *L*/in^2[Eq. (8-7)]
Excess At	2700 At/pole
Commutating-pole air gap	0.609 in [Eq. (8-8)]

8.8 MAGNETIC CALCULATIONS

Flux Paths. Figure 8-40 shows the paths of the main-pole flux for a typical medium-sized machine. The commutating poles and the compensating slots are not shown. Saturation calculations involve only half the length of a complete flux loop, because that is all that one field coil accommodates. Except for the main-pole air gap and the rotor teeth ampere-turns, the calculations are simple. They require (1) the determination of flux densities by dividing the flux in a section by its cross-sectional area, $\beta = \Phi$ area; (2) reading a magnetization curve for the material involved to find the ampereturns per inch needed for the density; and (3) finding the total ampere-turns for the part by multiplying the length of the portion of the path by those ampere-turns per inch. Typical magnetization curves are shown in Fig. 8-41.

The rotor core is usually built up of sheet steel laminations 0.017 to 0.025 in thick. Because of burrs and surface coatings, a stacking factor of 93% is common. The main poles use thicker laminations, and a factor of 95% is common. If the frame is also made up of laminations, a similar factor is necessary. Of course, a solid frame uses its full area.

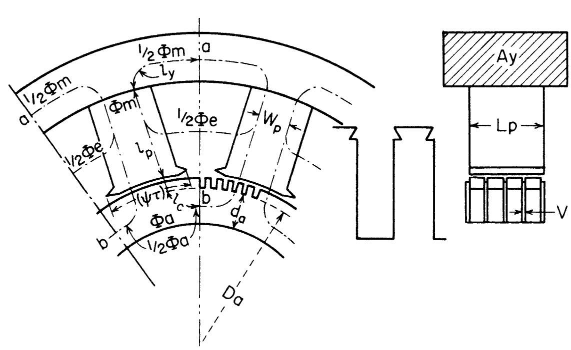


FIGURE 8-40 Paths of main and leakage fluxes.

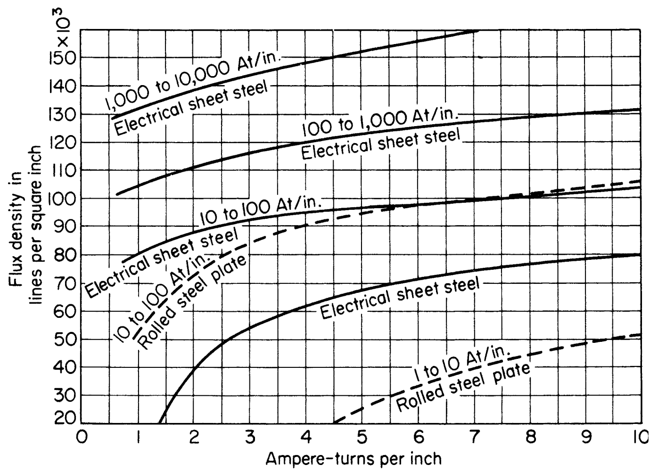


FIGURE 8-41 Magnetization curves.

The leakage flux ($^1\!/_2\ \phi_e$) in Fig. 8-40 from the main field coils must be included with the useful flux in the frame yoke and the pole body. Calculations depend on the actual machine dimensions and on the main field ampere-turns. However, the ampere-turns in these parts represent only a small part of the total required for the entire path, and it is usually accurate enough to estimate this leakage to be 12% of the useful flux normally and 20% at high saturations. For accurate calculations, the actual leakage can be plotted. No leakage fluxes are considered in computing the gap, teeth, or core densities.

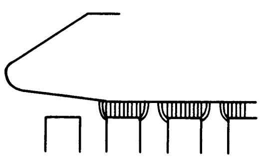


FIGURE 8-42 Distribution of flux in the air gap.

The Carter Coefficient and Gap Ampere-Turns. The presence of rotor slots, compensating slots, and vent ducts in the generator causes the actual densities in the main-pole air gap to be greater than for a smooth, solid core. Also, the average lengths of the flux paths are longer (see Fig. 8-42). The two effects may be lumped by assuming that the air gap is larger than measured mechanically. On considering the three factors (rotor slots, compensating slots, and vents) in succession, the formula

$$G_1 = G \times \frac{\text{G} + (\text{slot width}/5)}{G + (\text{slot width}/5)(1 - \text{slot width}/\text{slot pitch})} \tag{8-18}$$

gives the first corrected air gap G_1; this will closely approximate the effective air gap.

The ampere-turns across the gap will be

$$At_g = \beta_g \times 0.313 \times G_1 \tag{8-19}$$

The Rotor Teeth Ampere-Turns. For tooth densities below 100,000 lines/in^2, the ampere-turn drops in a tooth are so low that practically no flux will pass down the adjacent slot because the reluctance of air is so great. However, as tooth flux densities become larger, they produce very high ampere-turn drops from the top of the tooth to its bottom owing to saturation. Because these ampere-turns are also across the parallel flux path in the adjacent slot, when they are large enough, some useful flux will pass down the slot, relieve the tooth of some of its flux, and lower its actual density. If the tooth apparent density is calculated by assuming that all the flux across a slot pitch passes down the tooth, the actual density will be less than the apparent, depending on the amount of saturation.

The relation between the apparent tooth density β_{ta} and the actual tooth density β_t for different ratios of air area to iron area at any section of the tooth is shown in Fig 8-43. The *K* of these areas is

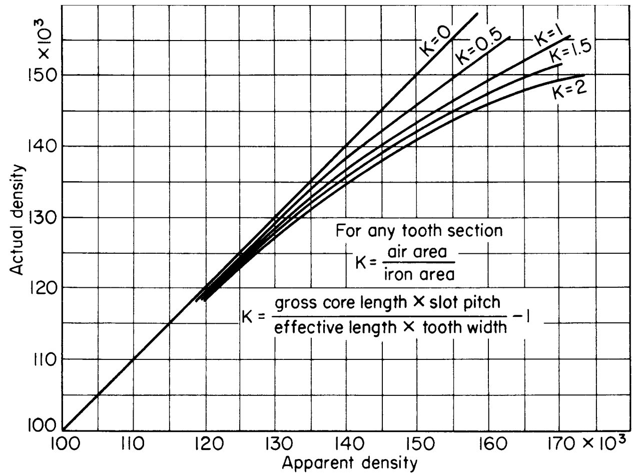


FIGURE 8-43 *K* curves.

$$K = \frac{\text{air area}}{\text{iron area}} = \frac{(\text{gross core length}) \times (\text{slot pitch})}{(\text{eff. core length}) \times (\text{tooth width})} - 1 \tag{8-20}$$

For accuracy in calculating tooth ampere-turns, it is desirable to divide the tooth into several parts, find the ampere-turns drop across each section, and total them. The flux density is found at the middle of each section, and the *K* ratio is calculated at the middle of each section.

Calculation of No-Load Saturation Data. Considering the 2500-kW, 700-V, 3571-A, 514-r/min generator, we have the values shown in Table 8-1. Using the magnetization curves of Fig. 8-41 and these data, the no-load saturation curve is calculated for several voltages. Note that 721 V is chosen in Table 8-2 on the assumption that the *IR* drop in the generator will not exceed 3%, or 21 V in this case. The generator (Fig. 8-44) must have this additional voltage induced in it for a 700-V terminal voltage. In the case of a motor, the induced voltage would be lower by the amount of the *IR* drop, or 679 V.

Full-Load Saturation Curve for a Compensated Machine. Figure 8-45 shows the calculated no-load saturation curve. For a well-compensated machine, the brushes will have little or no shift, and essentially no useful flux will be lost because of armature reaction. Only the armature-circuit-resistance *IR* drop need be considered, and the full-load excitation ampere-turns required can be read directly from the no-load saturation curve at the induced voltage.

For the 2500-kW generator, the excitation required at 721 V is 7520 At at full load.

TABLE 8-1 Magnetic Dimensions

Section	*K*	Net area, in^2	Eff. length, in
Frame yoke 6 × 17	—	102	13.85
Pole body 9 ½ × 15½	—	140	10.35
Compensating pole teeth (layout)	—	100	2.40
Effective air gap	—	—	0.268
Tooth 1 (upper ⅓)	0.92	10.8	0.75
Tooth 2 (middle ⅓)	0.96	10.3	0.75
Tooth 3 (bottom ⅓)	1.00	9.8	0.75
Core	—	79.3	7.15

Note: 1 in = 25.4 mm; 1 in^2 = 645 mm^2.

TABLE 8-2 Calculated Ampere-Turns per Pole

Volts	ϕ_t	Gap, $L = 0.268$	Tooth 1, $L = 0.75$, $K = 0.92$	Tooth 2, $L = 0.75$, $K = 2.96$	Tooth 3, $L = 0.75$, $K = 1.0$	Core, $L = 7.15$	Frame, $L = 13.85$	Pole, $L = 10.35$	C. tooth, $L = 2.40$	Total ampereturns
			Apparent β	Apparent β	Apparent β					
		β	Actual β	Actual β	Actual β	β	β	β	β	
			At/in	At/in	At/in	At/in	At/in	At/in	At/in	
		At	At	At	At	At	At	At	At	
			70,500	74,600	78,000	41,300	36,000	52,400	65,500	
420	65.5×10^6	33,800	70,500	74,600	78,000					
			5.4	7.2	9.0	2.1	6.5	2.8	4.2	
		2,850	5	5	5	15	90	30	10	3,018
			106,000	112,000	117,000					
630	98.3×10^6	50,800	106,000	111,300	116,200	62,000	54,000	78,600	98,300	
			130	210	320	3.7	11	9.2	6.2	
		4,280	100	160	240	25	150	95	15	5,065
			121,200	128,100	134,000					
721	112.5×10^6	58,200	120,500	126,500	132,300	71,000	61,800	90,000	112,500	
			440	660	1,000	5.7	14	26	225	
		4,900	330	495	750	40	195	270	540	7,520
			129,500	137,000	143,000					
770	120×10^6	62,100	128,000	134,600	139,000	75,800	66,000	96,000	120,000	
			710	1,200	1,850	7.9	16	46	410	
		5,230	535	900	1,390	55	220	475	985	9,790

Note: L = length of flux path, in; K = air area/iron area at particular position on tooth; apparent β = apparent flux density, lines/in^2; actual β = actual flux density, lines/in^2; At/in = ampere-turns per in; At = ampere-turns (1 in = 25.4 mm; 1 in^2 = 645 mm^2).

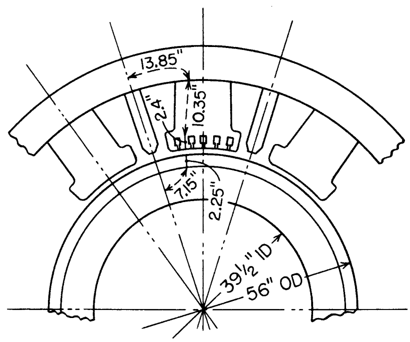


FIGURE 8-44 Cross section of a 2400-kW generator.

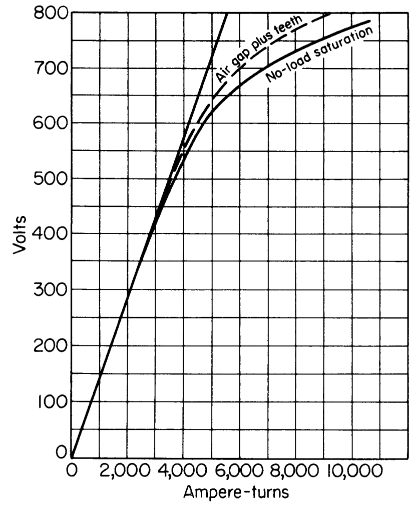


FIGURE 8-45 No-load saturation curves.

Full-Load Saturation Curve for a Noncompensated Machine. With commutating poles, there is no need for brush shift, but the uncompensated armature reaction will result in loss of useful flux as the load is increased. Figure 8-46 shows a method of calculating the additional ampere-turns excitation to replace this lost flux.

OBD = saturation curve of air gap plus teeth and pole face
BC = *IR* drop in armature circuit plus the brush drop. *B* = any point chosen on curve *OBD*
FB = *BE* = full-load-armature At/pole arc, or At/$p \times \psi$, laid off on a horizontal line

Through *E* and *F*, draw vertical lines of indefinite length. Move line *GI* vertically upward or downward parallel to *FBE* to a position *GHKI*, so that area *JGHOJ* area HABDIKH.

Through *B* draw a vertical line *BCK*. Then *HK* distortion ampere-turns for the load-current considered for point *B*.

Through *C*, draw a horizontal line of indefinite length cutting the no-load saturation curve at *A*.

CP = *HK*, to be extended from right at *C*
AP = total ampere-turns required at load current considered to maintain load at same value as at no load

By choosing several points, such as *B*, along the saturation curve and making the same calculations for each point, a full-load or any load saturation curve can be produced.

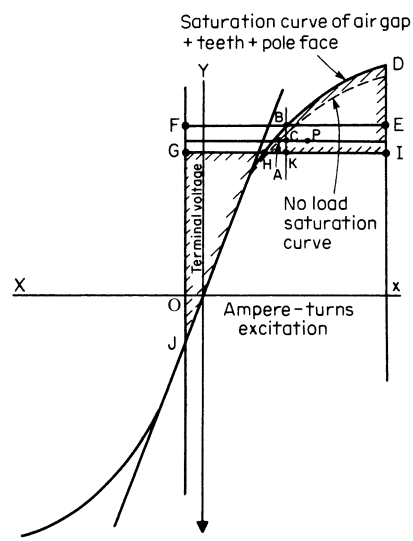


FIGURE 8-46 Calculation of load saturation curve.

Maximum Volts per Bar Calculations. The distorting ampere-turns resulting from imperfect compensation of the armature ampere-turns by the compensating winding are found by plotting the two and noting the maximum difference. This is done at the maximum-overload-current point.

The distortion factor (Figs. 8-32 and 8-45) is determined from the gap and teeth saturation curve (Fig. 8-45). At double load, the induced voltage is considered to be 740 V.

Volts between arms	700 V
No. of poles p	10
No. of commutating bars	375
Pole arc/pole pitch ψ	0.687
Distorting ampere-turns	1600 At
Distortion factor w/D of Fig. 8-32	1.06
Max. V/bar	$[(V \times p)/B] \times [w/(D \times \psi)]$ [Eq. (8-13)]
Max. V/bar	28.8 V/bar

This value is acceptable.

8.9 MAIN FIELDS

Main Field and Main-Field Heating. Figures 8-47 and 8-48 show three types of dc main fields. Small machines commonly use those of Fig. 8-47. They are wound on molds and then slipped on the poles. Type A is wound on an insulating spool, and type B uses an insulated steel spool for better heat transfer and mechanical protection.

The arrangement of Fig. 8-48 is common on large and medium-sized dc machines. The turns of the inner section are wound tightly on the insulated pole body to avoid air spaces between the pole and the coil. This permits maximum heat transfer. The second section is spaced away from the inner coil to permit the cooling air to flow over the maximum surface area possible. The thickness of a coil section is limited to about $1^1/_4$ to $1^3/_4$ in for a small temperature gradient within the coil.

All three types may use wire insulated with varnish, double cotton covering, or glass slivers in varnish. Air pockets which act as barriers to transfer of heat must be avoided, and so rectangular wire is common. Also, varnish or resin is liberally applied during winding or applied by vacuum impregnation after the coil is wound.

Design criteria suitable for all dc machines cannot be established, because the field cooling depends on air pressures from the armature rotation, the air-passage areas through the fields, and the radiation of heat from adjacent parts. These factors vary with machine design. However, on medium and large self-ventilated dc generators (built as in Fig. 8-48) empirical data are useful.

The main fields receive heat, not only from their own I^2R losses, but from heat radiated from the hot armature and the commutation coils. Also, the air cooling the coils is already heated by the

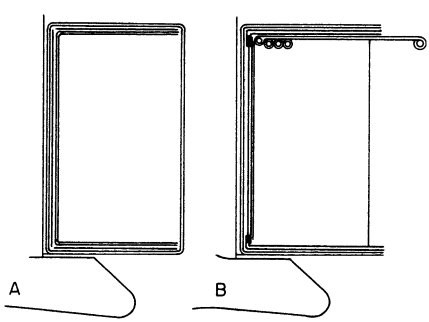


FIGURE 8-47 Two types of field-coil insulation, combined with fiber and metal spools, respectively.

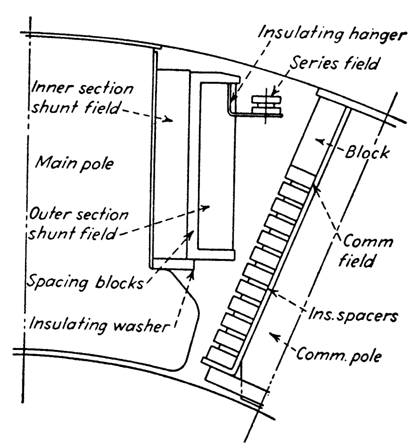


FIGURE 8-48 Ventilated field coils.

rotor. This lowers the temperature gradient for cooling the coils. The temperature rise of the fields must be calculated, not on the basis of the actual air temperature, but on the basis of the cool ambient-air temperature outside the machine. Figure 8-49 shows empirical data for such typical self-ventilated medium and large machines, built as shown in Fig. 8-48.

The "surface area" for these curves includes the entire periphery of the coil, because the heat transfer to the pole body is as effective as that to the air-cooled surfaces.

Little gain is made in cooling with increase in rotor velocities above 5000 ft/min because most of the armature air must pass through the limited field structure area. At high rotor speeds, the air is throttled owing to the high-velocity pressure drops.

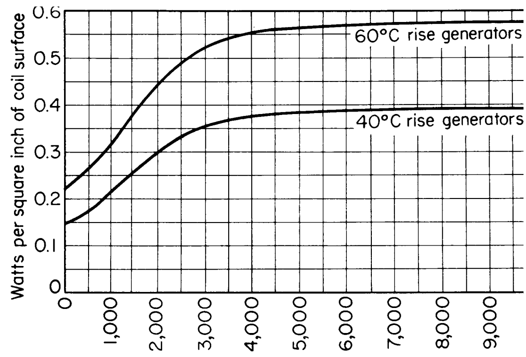


FIGURE 8-49 Main-field loss per surface area.

Main-Field Calculations. These are made by making a layout similar to that shown in Fig. 8-48. This permits the estimate of approximate mean length of turns ($\overline{L}_t$) for the sections.

The means of excitation and the particular application usually determine the *IR* drop of the main field. This is met in design by selection of the field wire cross-sectional area. This is calculated by Eq. (8-21).

$$\text{Conductor sectional area} = \frac{\text{At}/p \times \overline{L}_t \times p \times 8.25 \times 10^{-7}}{IR} \tag{8-21}$$

where At/p is the number of ampere-turns per pole needed, $\overline{L}_t$ is the mean length of turns, p is the number of coils in series, and IR is the required voltage drop.

Typical field calculations are

At/p	7520 At
Approx. $\overline{L}_t$	55 in
IR drop needed	90V
Conductor area [Eq. (8-21)]	0.038
Insulation conductor	0.018 in
Section of coil	6.78×1.6 in
Actual *IR*	86.5 V
Watts ($IR \times I$)	3380 W
W/in^2	0.362 W/in^2
W/in^2 allowed	0:388 W/in^2
Res. 75°C	2.21Ω
Coils in series	10
Copper	$0.162 \times 0.258 = 0.04$ in^2
Coil	24T high × 8lay. 192T/coil
Layout $\overline{L}_t$	55.65 in
$I = \text{A}t/t$	39.1 A
Surface $2H + tk(\overline{L}_t)p$	9350 in^2
Rotor velocity	7350 ft/min
Current density	977 A/in^2

These data indicate an acceptable field.

8.10 COOLING AND VENTILATION

Cause of Temperature Rise. The losses in a dc machine cause the temperature of the parts to rise until the difference in temperature between their surfaces and the cooling air is great enough to dissipate the heat generated.

Permissible measured temperature rises of the parts are limited by the maximum "hot-spot" temperature that the insulation can withstand and still have reasonable life. The maximum surface temperatures are fixed by the temperature gradient through the insulation from the hot spot to the surface.

The IEEE Insulation Standards have established the limiting hot-spot temperatures for systems of insulation. The *American National Standards Institute* Standard C50.4 for dc machines gives typical gradients for those systems, listing acceptable surface and average copper temperature rises above specified ambient-air temperatures for various machine enclosures and duty cycles. Typical values are 40°C for Class A systems, 60°C for Class B, and 80°C rise for Class F systems on armature coils. Class H systems usually contain silicones and are seldom used on medium and large dc machines. Silicone vapors can cause greatly accelerated brush wear at the commutator and severe sparking, particularly on enclosed machines.

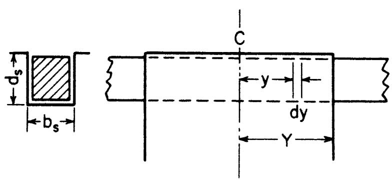


FIGURE 8-50 Heat paths in an armature conductor.

Temperature Gradients in Rotor Coils. Figure 8-50 represents a current-carrying conductor insulated from the core slot in which it is embedded. The hot spot is probably at the core centerline and near the center of the conductor. Heat will probably travel along the conductor to the end turn and also through the insulation to the iron. The amount of heat flowing in each direction is difficult to calculate. Also, variations in the coils, such as resin fill and tightness in the slots, make heat conductivity factors difficult to predict.

1. Assume that all the heat must travel down the conductor to the end turn. What will be the temperature difference in the conductor between the center of the core and its edge?

$$\text{Resistivity of copper at } 75°\text{C} = 8.25 \times 10^{-7}\Omega/\text{in}^3$$

$$\text{Thermal cond. copper} = 9.75 \text{ W/(in)(°C) for 1-in}^2 \text{ section}$$

Therefore, the energy crossing dy of Fig. 8-50 is

$$\text{Watts} = (I_c)^2 R_y = \frac{(I_c)^2(y)(8.25 \times 10^{-7})}{A} \tag{8-22}$$

where I_c is the conductor amperes, R_y the resistance of length y, and A the conductor cross-sectional area. The difference in temperature between two faces dy apart is

$$°\text{C} = \frac{(I_c^2)\,(y)(8.25 \times 10^{-7})}{A} \times \frac{dy}{A} \times \frac{1}{9.75} \tag{8-23}$$

and the difference in temperature between the center C and any point y is

$$°\text{C} = \frac{(I_c)^2(8.25 \times 10^{-7})}{A} \int_0^y \frac{y\,dy}{A} \times \frac{1}{9.75}$$

or

$$°\text{C} = 4.22\left(\frac{I_c}{A}\right)^2 (y)^2(10)^{-8} \tag{8-24}$$

Consider a current density of 2920 A/in^2 and a total core length of 16 in. Then the coil temperature gradient from the core center, with no ventilating ducts, to the edge is 28.8°C. This assumes that no heat passes through the insulation to the iron, and so medium and large machines normally use ventilating core ducts every few inches.

2. Assume that the end turns are so hot that no heat flows longitudinally down the coil. The I^2R loss of each inch of conductor length is

$$\text{Watts} = \frac{(I_c)^2(8.25 \times 10^7)}{A}$$

If the slot contains several conductors

$$\text{Watts} = (\text{ampere conductors})(\text{A/in}^2)(8.25 \times 10^{-7})$$

and the temperature difference between the bare conductor and the steel across the insulation is

$$°\text{C} = (\text{amp conductors})(\text{A/in}^2) \times \frac{\text{insulation thickness}}{2d_s + b_s} \times \frac{8.25 \times 10^{-7}}{0.003} \tag{8-25}$$

The factor 0.003 is the thermal conductivity of the insulation in watts per cubic inch per degree Celsius W/(in^3)(°C) difference. Thus, for 2142 ampere conductors per slot, 2920 A/in^2, a surface of two slot depths plus a slot width (times 1 in) = 5.07 in^2, and an insulation thickness of 0.051 in the temperature drop across the insulation is 17.65°C.

This figure cannot be considered precise because the thermal conductivity can vary widely with the insulation used and the presence of varying amounts of air in it. The conductivity figure for air is 0.0007, whereas that of mica is 0.007 W/(in^3)(°C). Also, heat moves along the coil. Because of these difficulties, empirical data from actual machines are more reliable and easier to use.

Heating of End Connections of Armature Windings. Small machines often have "solid" end windings banded down on insulated "shelf"-type coil supports. Larger machines are more heavily loaded per unit volume and usually have narrow coil supports, air spaces between the end turns, and ventilating air scouring both the top and bottom surfaces of the coil extensions.

With this construction, the approximate allowable product of ampere conductors per inch of outer circumference times the amperes per square inch for various rotor velocities is shown in Fig. 8-51 for a 40°C rise on the end turns.

Commutator Heating. A modern dc armature is shown in Fig. 8-52. The commutator diameter ranges from 55% to 85% of the rotor core, and the commutator necks joining the bars with the rotor

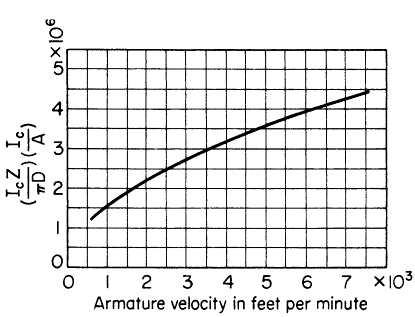


FIGURE 8-51 End-winding cooling.

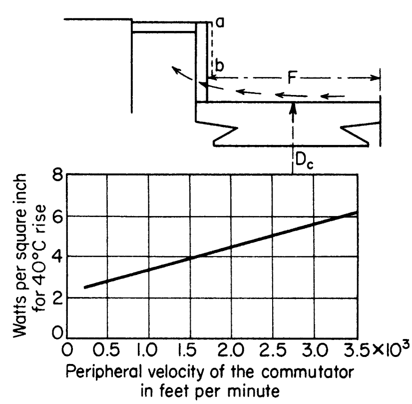


FIGURE 8-52 Temperature rise of a commutator.

winding extensions are usually separated from one another by air spaces, so that, when the armature revolves, air circulation is set up as shown by the arrows.

A typical relation between permissible watts per square inch of commutator surface and its peripheral velocity is shown in Fig. 8-52. The radiating surface is the commutator circumference times its face length. Neck area is not included.

The heat to be dissipated is that due to brush friction and the brush contact I^2R losses. There may be other losses due to poor commutation, brush chattering, and commutator surface, and, if so, the rise will be greater than indicated in Fig. 8-52. If commutation is very good and brush riding excellent, the temperature will be lower.

Application of Heating Constants. The paragraphs covering the design of the armature, main fields, compensating windings, and commutating windings included typical loading data such as ampere conductors per inch, amperes per square inch, flux densities, and watts per square inch of cooling surface. More accurate data depend on the exact arrangements used in a particular design. If possible, new design should be compared with similar machines which have already been tested. Any machine enclosure variation that restricts or increases the ventilation will affect the temperature rises.

8.11 LOSSES AND EFFICIENCY

Armature Copper I^2R Loss. At 75°C the resistivity of copper is $8.25 \times 10^{-7}\Omega/\text{in}^3$. Thus, for an armature winding of Z conductors, each with a length of $\overline{L}_t/2$ (half the mean length turn of the coil), each with a cross-sectional area of A and arranged in several parallel circuits, the resistance is

$$R_a = Z\frac{\overline{L}_t}{2A}\frac{8.25 \times 10^{-7}}{(\text{circuits})^2} \qquad \text{ohms} \tag{8-26}$$

The $\overline{L}_t$ is best found by layout, but an approximate value is

$$\overline{L}_t = 2[(1.35)(\text{pole pitch}) + (\text{rotor length}) \times 3] \tag{8-27}$$

There are also eddy current losses in the rotor coils, but these may be held to a minimum by conductor stranding in accordance with Eq. (8-17). Some allowance for these is included in the load loss.

Compensating, Commutating, and Series Field I^2R Losses. These fields also carry the line current, and the I^2R losses are easily found when the resistance of the coils is known. Their $\overline{L}_t$ is found from sketch layouts. At 75°C

$$R = T\frac{\overline{L}_t}{A}\frac{8.25 \times 10^{-7}}{(\text{circuits})^2}p \qquad \text{ohms} \tag{8-28}$$

where R is the field resistance in ohms, T the number of turns per coil, p the number of poles, $\overline{L}_t$ the mean length of turn, and A the area of the conductor. The total of these losses ranges from 60% to 100% of the armature I^2R for compensated machines and is less than 50% for noncompensated machines.

The brush I^2 loss is caused by the load current passing through the contact voltage drop between the brushes and the commutator. The contact drop is assumed to be 1 V.

$$\text{Brush } I^2R \text{ loss} = 2(\text{line amperes}) \qquad \text{watts} \tag{8-29}$$

Load Loss. The presence of load current in the armature conductors results in flux distortions around the slots, in the air gap, and at the pole faces. These cause losses in the conductors and iron that are difficult to calculate and measure. A standard value has been set at 1% of the machine output.

$$\text{Load loss} = 0.01(\text{machine output}) \tag{8-30}$$

Shunt Field Loss. Heating calculations are concerned only with the field copper I^2R loss. It is customary, however, to charge the machine with any rheostat losses in determining efficiency. Thus

$$\text{Shunt field and rheostat loss} = I_f V_{ex} \qquad \text{watts} \tag{8-31}$$

where I_f is the total field current and V_{ex} is the excitation voltage.

Core Loss. As seen from Fig. 8-53, the flux in any portion of the armature passes through $p/2$ c/r (cycles per revolution) or through $(p/2)[(\text{r/min})/60]$ Hz.

The *iron losses* consist of the *hysteresis loss*, which equals $K\beta^{1.6}fw$ watts, and the *eddy current loss,* which equals $K_e(\beta ft)^2w$ watts. K is the hysteresis constant of the iron used, K_e is a constant inversely proportional to the electrical resistance of the iron, β is the maximum flux density in lines per square inch, f is the frequency in hertz, w is the weight in pounds, and t is the thickness of the core laminations in inches.

The *eddy loss* is reduced by using iron with as high an electrical resistance as is feasible. Very high resistance iron has a tendency to have low flux permeability and to be mechanically brittle and expensive. It is seldom justified in dc machines. The loss is kept to an acceptable value by the use of thin core laminations, 0.017 to 0.025 in thickness.

Another significant loss is the *pole-face loss*. Figure 8-42 shows the distribution of flux in the air gap of a dc machine. As the armature rotates and the teeth move past the pole face, emfs are induced which tend to cause currents to flow across the pole face. These losses are included in the core loss.

Unfortunately, there are other losses in the core that may differ widely even on duplicate machines and that do not lend themselves to calculation. These include:

1. Loss due to filing of slots. When the laminations have been assembled, it will be found in some cases that the slots are rough and must be filed to avoid cutting the coil insulation. This burrs the laminations and tends to short circuit the interlaminar resistance.
2. Losses in the solid spider, core end plates, and coil supports from leakage fluxes may be appreciable.
3. Losses due to nonuniform distribution of flux in the rotor core are difficult to anticipate. In calculating core density, it is customary to assume uniform distribution over the core section. However, flux takes the path of least resistance and crowds behind the teeth until saturation forces it into the less used, longer paths below. As a result of the concentration, the core loss, which is about proportional to the square of the density, is greater than calculated.

Thus, it is not possible to predetermine the total core loss by the use of fundamental formulas. Consequently, core-loss calculations for new designs are usually based on the results from tests on similar machines built under the same conditions. Such test results are plotted in Fig. 8-54 for

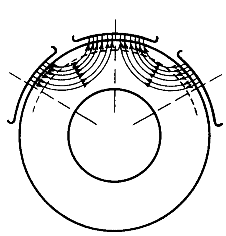


FIGURE 8-53 Distribution of flux in the armature.

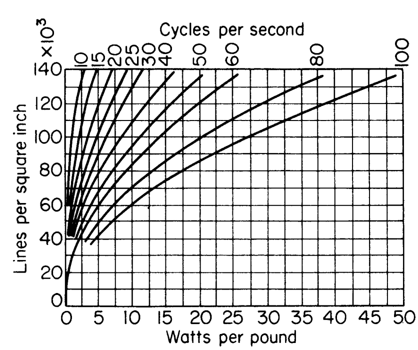


FIGURE 8-54 Iron-loss curves for a dc machine.

machines using ordinary laminations 0.017 in thick and a limited amount of filing. They do not include the pole-face losses, which would increase the values about 30%.

Brush Friction Loss. This loss varies with the condition of the commutator surface and the grade of carbon brush used. A typical machine has about 8-W loss/(in^2 of brush contact surface) (1000 ft/min) of peripheral speed when normal brush pressure of $2\frac{1}{2}$ lb/in^2 is used.

$$\text{Brush friction} = (8)\,(\text{contact area})\,\frac{\text{peripheral velocity}}{1000} \tag{8-32}$$

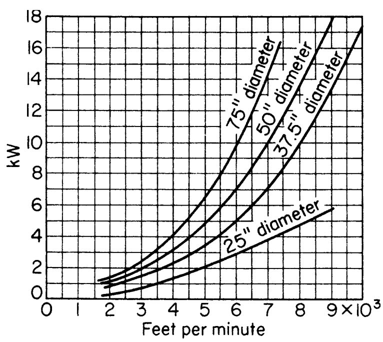


FIGURE 8-55 Friction and windage versus rotor velocity.

Friction and Windage. Most large dc machines use babbitt bearings and many small machines use ball or roller bearings, although both types of bearing may be used in machines of any size. The bearing friction losses depend on the speed, the bearing load, and the lubrication. The windage losses depend on the construction of the rotor, its peripheral velocity, and the machine restrictions to air movement. The two losses are lumped in most estimates because it is not practical to separate them during machine testing.

Figure 8-55 shows typical values of friction and windage losses for various rotor diameters referred to rotor velocities.

Efficiency. The efficiency of a generator is the ratio of the output to its input. The prime mover must supply the output and, in addition, the sum of all the losses. This is the input

$$\text{Efficiency} = \frac{\text{output}}{\text{input}} = \frac{\text{output}}{\text{output} + \text{losses}} \tag{8-33}$$

8.12 GENERATOR CHARACTERISTICS

The voltage regulation of a dc generator is the ratio of the difference between the voltage at no load and that at full load to the rated-load voltage. The characteristic is normally drooping as the load is increased, but it can rise because of series field effects or the action of circulating currents of communication at very low voltage operation.

For a dc generator, the terminal-voltage equation is

$$TV = E - IR = [K\phi_t)(\text{r/min}) - IR] \tag{8-34}$$

where E is the induced emf, IR is the armature circuit drop, K is a constant depending on the machine design, and ϕt is the total main-pole flux of the generator.

The regulation curves are easily calculated by using the no-load and full-load saturation curves shown in Fig. 8-56. The effect of the excitation method is found by the use of the field and rheostat IR line for self-excited machines and by the constant-ampere-turn line for separate excitation.

A separately excited compensated generator which is shunt-wound will have a voltage-load characteristic which will approach a straight line; it droops to full load an amount equal to the percent IR drop. There is little or no flux loss due to armature reaction or brush shift.

At voltages 10% or less of rated, the main-field strength is so weak that currents circulating in the coils short-circuited by the brushes at commutation may cause an increase in main-pole flux with load that causes a rising characteristic. These armature coils loop the main poles and their ampere-turns produce direct axis flux. A rising voltage characteristic can be undesirable, particularly if the generator supplies a dc motor whose speed is caused to rise with load, since this causes instability.

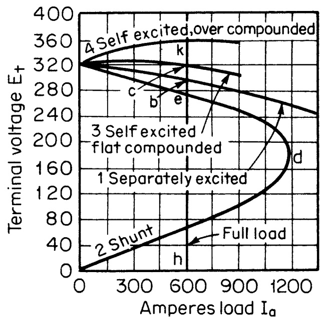


FIGURE 8-56 External characteristics versus excitation methods.

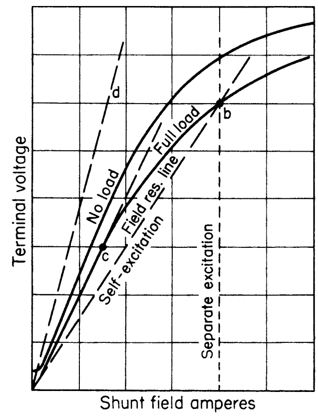


FIGURE 8-57 No-load and field-load saturation curves.

A separately excited noncompensated dc generator which is shunt-wound has a nonlinear loss of flux due to armature reaction as the load current is increased. It can be seen from Eq. (8-35) that this causes a characteristic which droops at an ever-increasing rate with load increase, giving a curve which is concave downward.

A self-excited noncompensated dc generator which is shunt-wound has its shunt-field excitation decreased as the terminal voltage drops. This results in a reduction of main-field ampere-turns and a loss of still more flux. This gives a severe droop which may be so great that, above a certain peak-load current, the terminal voltage will not be high enough to provide enough field current to maintain the voltage and load current and the voltage will collapse, as shown in *d* of Fig. 8-57.

Instability of Self-Excited Generators. A self-excited dc generator is unstable if the rheostat line does not make a definite intersection with the load-saturation curve (see Fig. 8-56). The shunt-field current is fixed by the terminal voltage, and the resistance is in the shunt-field circuit. Instability will exist if the slope of the rheostat line is nearly equal to or greater than the slope of a line tangent to the operating point on the saturation curve. In Fig. 8-57, point *b* is a stable operating condition, but point *c* is not, because a decrease in voltage decreases the shunt field ampereturns, and this produces a further decrease in voltage.

If the field circuit resistance were set at *d*, the self-excited generator would never build up beyond residual voltage. Another cause of failure to build up may be the connection of the shunt field. If the current flow due to residual voltage is such that it tends to kill the flux producing the residual voltage, no buildup occurs.

Compound-Wound DC Generators. The generators described above can be compounded by adding series fields excited by the load current. However, the resulting field strength of these fields is linear with load and the shape of the voltage-regulation curve is not changed thereby but is merely rotated upward or downward with the zero-load point as a pivot.

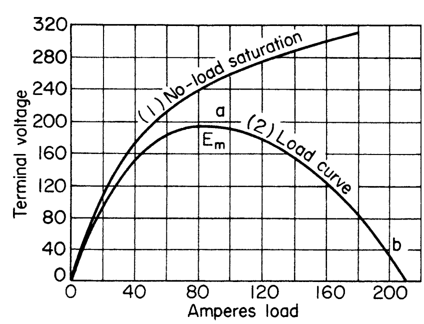


FIGURE 8-58 Characteristic curves of a series generator.

Series Generators. Curve 1 of Fig. 8-58 shows the relation between voltage and current if there is no armature resistance or armature reaction. This is actually the no-load curve of the machine obtained by separately exciting the series field. Curve 2 shows the actual relation between load current and terminal voltage. The total voltage drop is made up of a part caused by the decrease in flux by armature reaction and a part caused by the *IR* drop of the armature, brushes, and series fields.

Field Time Constants. The major delay in change of output voltage by an excitation change is caused by the inductance of the main fields. The time constant of the shunt field is the ratio of its inductance in henries to its resistance in ohms, and this ratio represents the time in seconds required for 63% of a field current change to occur when the excitation voltage is suddenly changed. In the case of the 2500-kW generator, a mean main-field inductance over the voltage range from zero to rated is 6.20 H. The main-field resistance is 2.21Ω. The field time constant is therefore 2.8 s.

The inductance L of a coil is the incremental change of flux linkages per incremental change in field current times 10^8. This is proportional to the slope of the saturation curve and is constant over the air-gap line. It is therefore a decreasing variable after the curve leaves the air-gap line (see Fig. 8-45). The overall inductance, as the voltage builds up from zero, is not so high as that of the air-gap portion or as low as at the rated-voltage point. A common compromise is the slope of a straight line drawn from zero voltage through the full-load point at rated voltage. For the 2500-kW generator the total flux at this point is 112.5×10^6 lines. With a leakage flux of 12%, each coil has a flux of 12.6×10^6 lines (see Table 8-2). As indicated earlier [see Eq. (8-21) and surrounding text], each coil has 192 turns and there are 10 coils in series. The field current is 39.1 A.

$$L = \frac{\phi T}{I_f} \times 10^{-8} = \frac{(12.6 \times 10^6)(192)(10)}{39.1} \times 10^{-8} = 6.2\ \mathrm{H} \qquad (8\text{-}35)$$

$$\text{Time constant} = \frac{L}{R} = 6.2/2.21 = 2.8\ \mathrm{s} \qquad (8\text{-}36)$$

This value is typical for large machines. Smaller generators have less copper in their fields and lower time constants. In cases where drive systems must have very rapid voltage adjustments, it is common to provide large forcing voltages on the field to overcome the inductive lag. These sudden excitation changes may be 4 to 10 times the *IR* drop of the field. This effectively reduces the time constant to one-fourth or one-tenth its normal value.

Armature-Circuit Time Constants. Compensating windings effectively lower the inductances of the armature circuit. The 2500-kW generator developed in this section has an armature-circuit inductance of 0.0001929 H and a circuit resistance of 0.00398Ω for a time constant of 0.048 s. This value is typical for large dc machines. Smaller noncompensated units have longer time constants.

8.13 TESTING

Factory Tests. These depend on the size, application, and design of the dc generator. The American National Standards Institute (ANSI) C50.4 for dc machines includes lists of recommended tests for dc generators and motors. The IEEE Test Code for dc machines covers recommended methods to be used for these tests.

8.14 GENERATOR OPERATION AND MAINTENANCE

General. Despite its rugged construction, a dc machine is a delicate device. Factory tests on large units may cost thousands of dollars and must be performed carefully to adjust the generator to obtain the best possible characteristics and commutation. Owing to shipping requirements, the generator may then have to be disassembled and shipped in several pieces. If the final assembly is not correctly accomplished, not only have the factory tests been wasted but the machine may be damaged.

The manufacturer's instruction book should be studied carefully.

Before Installation. Upon arrival, the generator should be inspected for damage and to be sure it is dry. If it is wet, consult the manufacturer. Drying out with heat should be done only by slowly raising the generator temperature to 100°C so that moisture can escape without forming gas pockets within the insulation.

If the generator is dry and clean, the windings should be checked with a megger for insulation resistance to ground measurements. If any readings less than 1 MΩ are found, check with the manufacturer.

Alignment. After the machines are installed and grouted to the foundations, all couplings should be opened and alignments of all shafts finally checked. Regardless of whether solid or flexible couplings are used, the alignment should be as accurate as possible. The difference between the bottom and the top openings should not exceed 0.002 in for 12 in of flange diameter, and the large opening should be at the top. Regardless of the size of coupling, the difference should not exceed 0.004 in. Differences at the side should not exceed 0.001 in. Shafts should be rotated 180° and rechecked.

The frame should be set on the *magnetic center* of the core. This position can be located by setting the armature in rotation and forcing it to oscillate longitudinally the full end play of the bearing by pushing on the end of the shaft. While the rotor is coasting and oscillating freely, excite the main field. The stator can then be shifted so that the rotor position with excitation coincides with the center of bearing end play.

Air gaps between the rotor and poles should be uniform. A typical limit of variation is 0.010 in. The brushes should ride properly on the commutator surface at both extremes of bearing end play.

Prerunning Checks. The circumferential position of the brushes on the commutator is important for commutation and also to provide the voltage characteristics set at the factory. Brushes should be on the factory test setting. The toes of the brushes should be aligned and should have no skew. The spacing between adjacent arms of brushes should be identical within 0.032 in. The brushes should move freely in their holders and should have a pressure against the commutator of 2 to 3 lb/in^2 on the basis of brush cross section. The faces of the brushes should accurately match the curvature of the commutator surface.

The polarity of the main fields may be checked by tracing the wiring around the frame or by lightly exciting the fields and using a compass around the frame behind the poles.

The oiling system for the bearings should be checked and the oil rings tested for freedom. The entire machine, particularly its air gaps, should be inspected for foreign material.

Running Checks. Note any unusual noise as the unit is brought up to speed. Bearing temperatures should level out at acceptable values within a few hours.

The voltage should be slowly raised at no load and commutation observed. If satisfactory, the voltage should be raised to 110% of rated and then reduced.

The generator may then be loaded gradually while commutation is observed, until rated current is reached. If commutation remains satisfactory until stable temperatures are achieved, the generator is ready for work.

Shunt-Wound Generators in Parallel. *A* and *B* of Fig. 8-59 are two similar generators feeding the same bus bars *C* and *D*. If *A* tends to take more than its share of the total load, its voltage falls and more load is automatically thrown on *B*. Also, if the driver of one of the generators slows down to stop, the emf of the machine falls until the other generator starts to drive it as a motor. This continues until its driver takes over again.

The external characteristics of the two machines are shown in Fig. 8-60. At voltage *E*, the currents in the generators are I_a and I_b, and the line current is $I_a + I_b$. To make machine *A* take more of the load, its excitation must be increased to raise its characteristic curve. If a 1000-kW generator and a 500-kW machine have the same regulation curves, the machines will divide the load according to their respective capacities, as shown in Fig. 8-61.

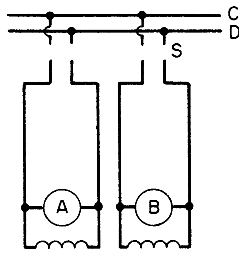


FIGURE 8-59 Shunt generators in parallel.

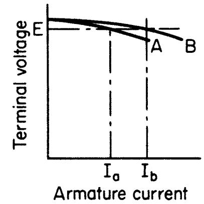


FIGURE 8-60 External characteristics of two shunt-wound generators in parallel.

Compound-Wound Generators in Parallel. *A* and *B* of Fig. 8-62 are two compound-wound machines. If *A* tends to take more than its share of the load, the series excitation of *A* increases, its voltage rises, and it takes still more of the load. Thus, the operation is unstable. If this continues until *A* takes all the load and the voltage of *B* drops to the point that *A* reverses the current in *B*, *B* will be driven as a motor. With the reversed current in the series field of *B* it becomes a differentially compounded motor, and the series weakens the flux to speed up the motor. This may progress to a point at which the unit may be damaged mechanically and electrically.

To prevent this, a bus bar of large section and of negligible resistance, called an *equalizer bus*, is connected from *e* to *f* (Fig. 8-62). Points *e* and *f* are then practically at the same potential. Therefore, the current in each series coil is independent of the current in its particular generator, is inversely proportional to the resistance of the coils, and is always in the same direction.

When a single compound generator has too much compounding, a shunt in parallel with the series field coils will reduce the current in these coils and so reduce the compounding. When compounded generators are operating in parallel using an equalizer bus, the current in the series field coils depends only on the resistance of the coils and a shunt connected across one of them is actually across all of them, reducing the compounding of all but not disturbing the relative compounding between the machines. To reduce the compounding of a single machine, it is necessary to place a resistance in series with the coils. This may require a large resistor to handle the large load current it must carry.

Maintenance. Except for the commutator and its brushes, maintenance of dc machines differs little from that of other rotating electrical machines. Proper lubrication must be provided for the bearings, and the machine must be kept clean and dry. In addition, the brushes should be checked periodically for commutation, riding ability, freedom of motion in the holders, pressure, and length.

Because the commutator necks are not insulated and receive full voltage, conducting dust from brush wear or from ventilating air can cause creepage currents between the risers and ground over insulated surfaces. To avoid this, the dc generator must be cleaned and blown out with clean, dry air at regular intervals. Air pressures above 25 lb/in^2 should not be used because of the danger of lifting the edges of insulating tape. The effectiveness of the cleaning program should be verified occasionally by megger readings.

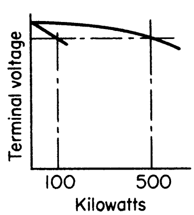


FIGURE 8-61 Division of load between two shunt generators in parallel.

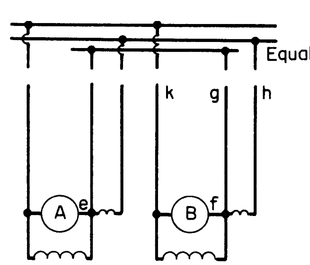


FIGURE 8-62 Compound generators in parallel.

Poor Commutation. Sparking and bar burning are usually due to one or more of the following causes:

1. *Brushes not in the proper position.*
2. *Incorrect spacing of brushes*. This may be checked by marking an adding-machine tape around the commutator.
3. *Projecting-bar-edge mica*. Mica between bars should be undercut about 0.063 in below the commutating surface, but occasionally slivers of mica are left inadvertently along the bar.
4. *Rough or burned commutator*. The commutator should be ground according to the manufacturer's instruction book.
5. *Grooved commutator*. This may be prevented by properly staggering the brush sets so that the spaces between the brushes of an arm are covered by brushes of the same polarity of other arms.
6. *Poor brush contact.* This is due to improper fitting of the brushes to the commutator surface. To seat the brushes, sandpaper should be moved between the commutator and the brush face. Emery cloth should not be used because its abrasive is conducting.
7. *Worn brushes replaced by others of wrong size or grade.*
8. *Sticking brushes.* These brushes do not move freely in their holders so that they can follow the irregularities of the commutator.
9. *Chattering of the brushes*. This is usually due to operation at current densities below 35 A/in^2 and must be corrected by lifting brushes to raise the density or by using a special grade of brush.
10. *Vibration*. This may be due to poor line up, inadequate foundations, or poor balance of the rotor.
11. *Short-circuited turns on the commutating or compensating fields*. These may be obvious on inspection but usually must be found by passing ac current through them for voltage-drop comparisons.
12. *Open or very high resistance joints between the commutator neck and the coil leads.* In this case, the bar at the bad joint will usually be burned.
13. *An open armature coil*. A broken coil conductor produces an effect similar to that produced by the poor joints described in the previous item. For emergency operation, the open coil may be opened at both ends, insulated from the circuit, and a jumper placed across the two affected necks. Since some sparking will probably result, operation should be limited.
14. *Short-circuited main-field coils*. With the resulting unbalanced air-gap fluxes under the poles, large circulating currents must be expected even with good armature cross connections. The offending coil may be found by comparing voltage drops across the individual coils.
15. *Reversed main-field coil*. This is an extreme case of the one described in the previous item.
16. *Overloading*.

8.15 SPECIAL GENERATORS

General. The adaptability of the dc generator for specific uses has led to the development of many special generators. These machines over the years made a significant contribution to industrial progress. However, most of these special applications have disappeared or are now being met with other devices such as silicon controlled rectifiers or programmed control of field currents to the main dc generator.

Synchronous Converters. Of all the special generators, this was one of the earlier and most widely used. It was the principal dc power source for streetcars and interurban lines. It was a most ingenious device, combining in a single armature and winding an ac motor taking its current from the lines through slip rings at the rear and a dc generator providing dc power from a commutator on the front

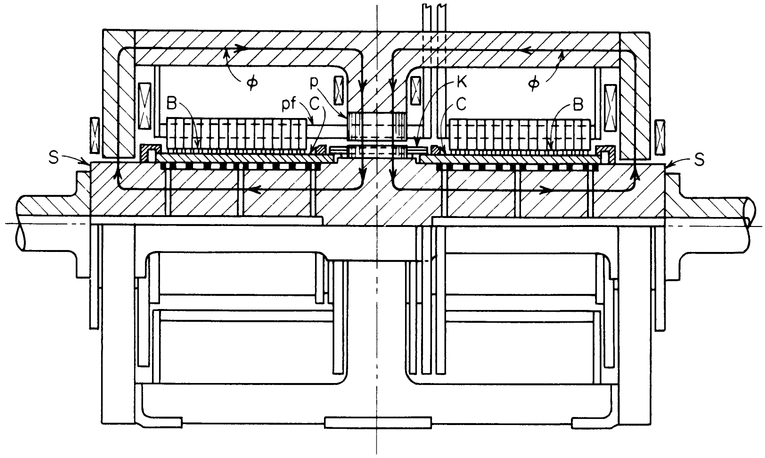


FIGURE 8-63 Brush-type homopolar generator.

end. Because the flow of the currents was in opposition, the resulting rotor winding could be small in cross section. A single stator provided flux for both functions. With the decline of street railway systems, the synchronous converter disappeared.

Rotating Regulators. These dc machines had trade names like Rototrol, Regulex, and Amplidyne. They, too, have been replaced by solid-state devices. In addition to having fields for feedback intelligence, response was enhanced using self-excited shunt fields tuned to the air-gap line or by means of cross-magnetization from armature reaction.

Three-Wire Devices. Because three-wire dc circuits are no longer in use, balancer sets and three-wire generators are relics in school labs or museums.

Homopolar or Acyclic DC Generators. The single-pole machine principle still fascinates electrical engineers and several research and development labs continue to study new arrangements of its basic parts. Fundamentally, it consists of a single conductor moving through a uniform single-direction flux with a collector at each end of the conductor. The output is a steady ripple-free pure dc current and no commutation. Currents reaching 270,000 A at 8 V were provided by one commercial unit shown in Fig. 8-63. Recent efforts have been mainly to use liquid metals to take the large currents from the rotating collectors and to obtain higher voltages by connecting units in series. Some success has been possible, but restricting the sodium potassium to the collector area has proved difficult.

BIBLIOGRAPHY

Alerich, W. N., *Electricity 3: DC Motors & Generators, Controls, Transformers,* Albany, N.Y., Delmar, 1981.

Alerich, W. N., and Keljik, Jeff, *Electricity 4: DC Motors & Generators*, Albany, N.Y., Delmar, 2001.

Blalock, G. C., *Direct-Current Machinery,* New York, McGraw-Hill, 1947.

Chapman, S. J., *Electric Machinery Fundamentals*, New York, McGraw-Hill, 1998.

Clayton, A. E., *The Performance and Design of Direct Current Machines, a Textbook for Students at Universities and Technical Schools,* London, Pitman, 1947.

Clayton, A. E., *The Performance and Design of Direct Current Machines,* London, Pitman, 1959.

Elliott, T. C., Chen, Kao, and Swanekamp, Robert, New York, McGraw-Hill, 1998.

Heller, S., *Direct Current Motors and Generators: Repairing, Rewinding, and Redesigning,* New Canaan, Conn., Datarule, 1982.

Kloeffer, R. G., Brenneman, J. L., and Kerchner, R.M., *Direct Current Machinery,* New York, Macmillan, 1950.

Langsdorf, A. S., *Principles of Direct-Current Machines,* New York–London, McGraw-Hill, 1940.

Lister, E. C., and Rusch, R. J., *Electric Circuits and Machines*, New York, McGraw-Hill, 1993.

Liwschitz-Garik, M., *Direct-Current Machines,* Princeton, N.J., Van Nostrand, 1956.

Rieger, K., *D-C Generators and Motors*. Scranton, Pa., International Correspondence Schools, 1968.

Siskind, C. S., *Direct-Current Machinery,* New York, McGraw-Hill, 1952.

Young, E. L., *D-C Machines,* Scranton, Pa., International Correspondence Schools, 1975. (Based on material provided by Scott Hancock; rev. by E. L. Young.)

SECTION 9
HYDROELECTRIC POWER GENERATION

U.S. Army Corps of Engineers
Hydroelectric Design Center

CONTENTS

9.1 GENERAL

9.1.1 Introduction

Hydropower is produced when kinetic energy in flowing water is converted into electricity. Hydropower has been a significant source of electrical energy in the United States since the early 1900s when manufacturers recognized and harnessed its tremendous potential to develop and build entire industries. Traditionally, hydropower has been a low-cost, reliable energy source. It utilizes a renewable fuel (water) that can be sustained indefinitely, and is free of fossil fuel emissions. And because hydroelectric generators are especially suited for providing peaking power, hydropower complements thermal generation and improves overall power production efficiency. Hydroelectricity presently constitutes approximately 10 percent of the United States' energy supply, which is enough to meet the needs of 28.3 million consumers.

9.1.2 Notations

a = celerity or speed of sound in water, feet/second

BOD = biological oxygen demand, parts per million/day

D = Winter-Kennedy piezometric pressure differential, feet

DO = dissolved oxygen, parts per million

E = specific energy, foot-pounds (force)/pound (force)

E_{rel} = relative efficiency, kilowatts/feet$^{1/2}$

E_t = turbine efficiency, percent or decimal

$E_{t\text{-}g}$ = combined turbine-generator efficiency, percent or decimal

G = local acceleration of gravity, feet/second^2

H = total net head or total dynamic head, feet

H_b = barometric pressure head, feet

H_d = design head (head of best efficiency), feet

HP = turbine output, horsepower

H_0 = initial piezometric head, feet

K = radius of gyration, feet

kW = generator output, kilowatts

L = length of water conduit, feet

MW = generator output, megawatts

MVA = generator or transformer capacity, megavolts-amperes

MVAR = generator output, reactive, megavars

N = rotational speed, revolutions/minute

N_s = specific speed, revolutions/minute-horsepower$^{1/2}$/head$^{5/4}$

Q = volumetric flow rate, feet^3/second

Q_{20} = 20 percent flow exceedence (time flow value is exceeded), percent

Q_{30} = 30 percent flow exceedence (time flow value is exceeded), percent

T or t = time, seconds

V = flow velocity, feet/second

V_0 = initial flow velocity, feet/second

W = weight, pounds (force)

WK^2 = angular inertia, pound-feet^2

γ = specific weight of water, pounds/foot^3

9.1.3 Nomenclature

The following terms are commonly used to describe hydroelectric equipment, facilities, and production:

Afterbay (tailrace). The body of water immediately downstream from a power plant or pumping plant.

Appurtenant structures. Intakes, outlet works, spillways, bridges, drain systems, tunnels, towers, etc.

Auxiliary power. The electric system supply to motors and other auxiliary electrical equipment required for operation of a generating station.

Base loading. Running water through a power plant at a roughly steady rate, thereby producing power at a steady rate.

Base load plant. Powerplant normally operated to take all or part of the minimum load of a system, and which consequently runs continuously and produces electricity at an essentially constant rate. Operated to maximize system mechanical and thermal efficiency and minimize operating costs.

Bulkhead. A one-piece fabricated steel unit that is lowered into guides and seals against a frame to close a water passage in a dam, conduit, spillway, etc.

Bulkhead gate. A gate used either for temporary closure of a channel or conduit before dewatering it for inspection or maintenance or for closure against flowing water. Bulkhead gates nearly always operate under balanced pressures.

Cavitation damage. Pitting and wear damage to solid surfaces (e.g., the blades of a hydraulic turbine) caused by the implosion of bubbles of water vapor in fast-flowing water.

Cofferdam. A temporary barrier, usually an earthen dike, constructed around a worksite in a reservoir or on a stream. The cofferdam allows the worksite to be dewatered so that construction can proceed under dry conditions.

Crest. The top surface of a dam or high point of a spillway control section.

Dam. A concrete and/or earthen barrier constructed across a river and designed to control water flow or create a reservoir.

Dewater (unwater). To drain the water passages and expose the turbine runner. Generally requires closing of an isolation valve or lowering of the headgates, and opening of the penstock drain valves.

Draft tube. Part of the powerhouse structure designed to carry the water away from the turbine runner.

Fish bypass system. A system for intercepting and moving fish around a dam as they travel downriver toward the ocean.

Fish ladders. A series of ascending pools constructed to enable salmon or other fish to swim upstream around or over a dam.

Fish screen. A screen across the turbine intake of a dam, designed to divert the fish into a bypass system.

Fish passage facilities. Features of a dam that enable fish to move around, through, or over without harm. Generally an upstream fish ladder or a downstream bypass system.

Forebay (headrace). The body of water immediately upstream from a dam or hydroelectric plant intake structure.

Generator. The machine that converts mechanical energy into electrical energy.

Head. The difference in elevation between two specified points, for example, the vertical height of water in a reservoir above the turbine.

High-head plant. A powerplant with a head over 800 ft.

Hydraulic losses. Energy loss in water passages primarily due to velocity losses at trash racks, intakes, transitions, and bends, and friction losses in pipes.

Intake. The entrance to a conduit through a dam or a water conveyance facility.

Intake structure. The concrete portion of an outlet works including trashracks and/or fish screens, upstream from the tunnel or conduit portions. The entrance to an outlet works.

Low-head plant. A powerplant with a head less than 100 ft.

Medium-head plant. A powerplant with a head between 100 and 800 ft.

Multipurpose project. A project designed for two or more water-use purposes. For example, any combination of power generation, irrigation, flood control, municipal and/or industrial water supply, navigation, recreation, and fish and wildlife enhancement.

Operating rule curve. A curve, or family of curves, indicating how a reservoir is to be operated under specific conditions and for specific purposes.

Outlet works. A combination of structures and equipment located in a dam through which controlled releases from the reservoir are made.

Peaking plant. A powerplant in which the electrical production capacity is used to meet peak energy demands. The site must be developed to provide storage of the water supply and such that the volume of water discharged through the units can be changed readily.

Penstock. A pipeline or conduit used to convey water under pressure from the supply source to the turbine(s) of a hydroelectric plant.

Pool. A reach of stream that is characterized by deep, low velocity water and a smooth surface.

Powerhouse. Primary structure of a hydroelectric dam containing turbines, generators, and auxiliary equipment.

Pumped storage plant. Powerplant designed to generate electric energy for peak load use by pumping water from a lower reservoir to a higher reservoir during periods of low energy demand using inexpensive power, and then releasing the stored water to produce power during peak demand periods.

Reservoir. A body of water impounded in an artificial lake behind a dam.

Runoff. Water that flows over the ground and reaches a stream as a result of rainfall or snowmelt.

Run-of-the-river plant. A hydroelectric powerplant that operates using the flow of a stream as it occurs and having little or no reservoir capacity for storage or regulation.

Single-purpose project. A project in which the water is used for only one purpose, such as irrigation, municipal water, or electricity production.

Spill. Water passed over a spillway without going through turbines to produce electricity. Spill can be forced, when there is no storage capability and stream flows exceed turbine capacity, or planned, for example, when water is spilled to enhance downstream fish passage.

Spillway. The channel or passageway around or over a dam that passes normal and/or flood flows in a manner that protects the structural integrity of the dam.

Standby power. Frequently provided as a backup for operating gates and valves in the event the principal power supply (usually electrical) fails. Includes engine-driven-generators or hydraulic oil pumps, each of which could be powered by gasoline, diesel, or propane, and power takeoffs on trucks or tractors. On small-sized gates or valves, the standby power is often hand-operated, such as a hand pump or crank.

Stoplogs. Large logs, planks, steel or concrete beams placed on top of each other with their ends held in guides between walls or piers to close an opening in a dam, conduit, spillway, etc., to the passage of water. Used to provide a cheaper or more easily handled means of temporary closure than a bulkhead gate.

Storage reservoir. A reservoir having the capacity to collect and hold water from spring time snowmelts. Retained water is released as necessary for multiple uses such as power production, fish passage, irrigation, and navigation.

Surge tank. A large tank, connected to the penstock, used to prevent excessive pressure rises and drops during sudden load changes in plants with long penstocks.

Switchyard. An outdoor facility comprised of transformers, circuit breakers, disconnect switches, and other equipment necessary to connect the generating station to the electric power system.

Tailrace. See Afterbay.

Tailwater. The water in the natural stream immediately downstream from a dam.

Transformer. An electromagnetic device used to change the magnitude of voltage or current of alternating current electricity or to electrically isolate a portion of a circuit.

Trashrack. A metal or reinforced concrete structure placed at the intake of a conduit, pipe, or tunnel that prevents large debris from entering the intake.

Trashrake (trash rake). A device that is used to remove debris, which is collected on a trashrack to prevent blocking the associated intake.

Turbine, hydraulic. An enclosed, rotary-type prime mover in which mechanical energy is produced by the force of water directed against blades or buckets fastened in an array around a vertical or horizontal shaft.

Turbine runner (water wheel). The rotor-blade assembly portion of the hydraulic turbine where moving water acts on the blades to spin them and impart energy to the rotor.

Unwater. See Dewater.

Wicket gates. Adjustable gates that pivot open around the periphery of a hydraulic turbine to control the amount of water admitted to the turbine.

9.2 HYDROELECTRIC POWERPLANTS

To determine the optimal location, size, and layout of a hydroelectric powerplant, numerous factors must be considered including the local topography and geologic conditions, the amount of water and head available, power demand, accessibility to the site, and environmental concerns. The overriding consideration in the design of a hydroelectric powerplant is that it adequately perform its function and is structurally safe.

9.2.1 Principal Features

The principal features of a hydroelectric facility are the dam, reservoir, spillway, outlet works, penstocks, powerhouse, fish passage facilities (if fish protection is required), surge tanks, and switchyard. Most hydroelectric powerplants are located at or immediately adjacent to a dam. Some plants, however, are located away from the dam, such as at the lower end of a pressure penstock, power tunnel, or power canal, or at a drop in an irrigation canal. In general, a powerplant is situated so that the penstocks will be as short as practicable in order to minimize the cost of the penstocks and the associated hydraulic losses, and to avoid the necessity for surge tanks.

Hydropower developments can be classified as either low-, medium-, or high-head projects. Figure 9-1 shows in outline the most common arrangements, and illustrates some of the features listed in the Sec. 9.13 for the various developments. Other sources of hydropower involve the use of ocean waves or tidal changes to generate electricity. These technologies are not as well developed as the more conventional hydropower sources and are not covered in this chapter.

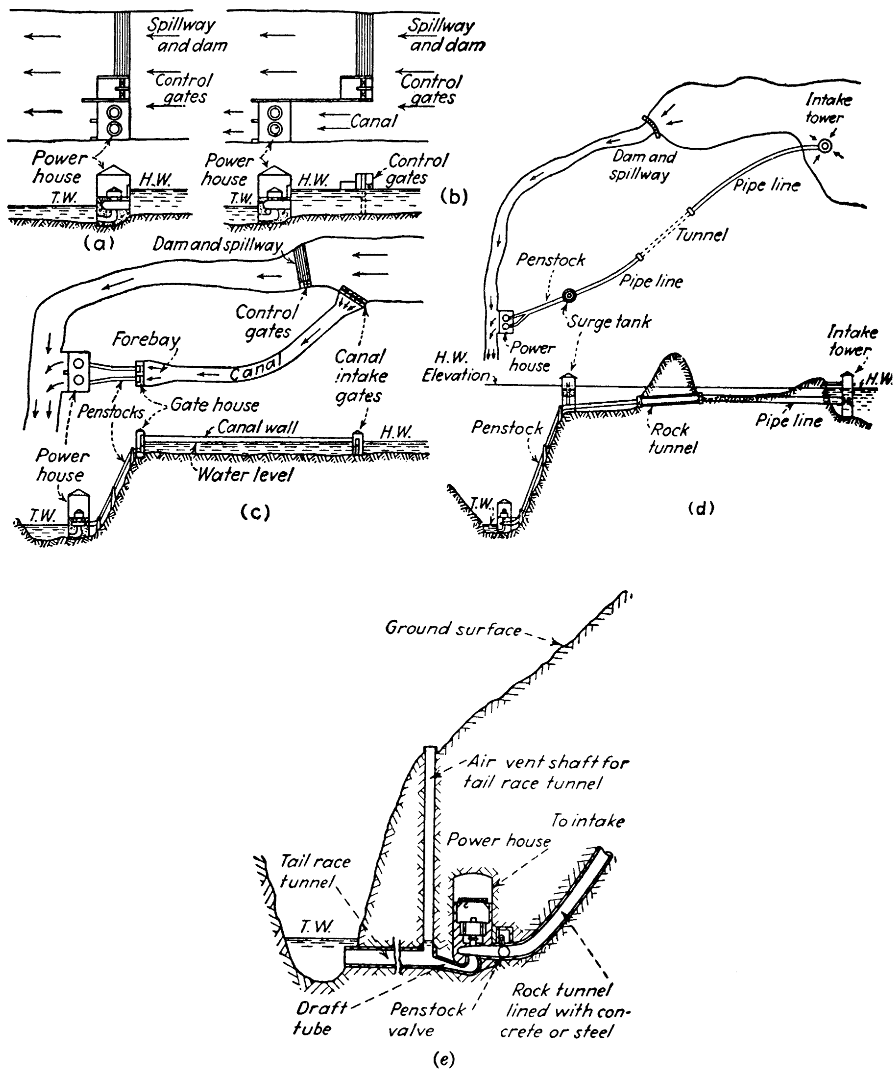


FIGURE 9-1 Outline sketches of several typical hydropower developments: (*a*) low-head development with dam, spillway, and powerhouse as an integral unit; (*b*) low-head development with a short intake canal and powerhouse separate from the dam; (*c*) medium-head development with a long intake canal, gatehouse, and penstocks connecting the forebay with the powerhouse; (*d*) high-head development with a large storage reservoir, pipeline, and tunnel leading to a surge tank at the upper end of the penstocks—powerhouse at the lower end of the penstocks is a considerable distance from the dam and spillway; (*e*) outline sketch of underground powerplant, showing penstock and tailrace tunnels.

9.2.2 Powerhouse Structure

The powerhouse foundation and superstructure contain the hydraulic turbine, water passages including draft tube, passageways for access to the turbine casing and draft tube, and sometimes the penstock valve. The superstructure also typically houses the generator, exciter, governor system, station service, communication and control apparatus, and protective devices for plant equipment and related auxiliaries as well as the service bay, repair shop, control room, and offices. The transformers and switchyard are usually located outdoors adjacent to the powerhouse and are not an integral part of it. Cranes are provided in the powerhouse to handle the heaviest pieces of turbine and generator and sometimes extend over the penstock valves. Alternative powerhouse designs have included separate cranes for the penstock valves. Another common powerhouse design is the outdoor type where the operating floor is placed adjacent to the turbine pits with the generator located outdoors on the roof of a one-story structure. In the outdoor type, each generator is protected by a light steel housing, which is removed by the outdoor gantry crane when access to the machine is necessary for other than routine maintenance. The erection and repair space is in the substructure and has a roof hatch for equipment access. The outdoor design reduces initial construction costs of the powerplant. However, the choice of indoor, semi-outdoor, or outdoor type is dictated not only by consideration of the initial cost of the structure with all equipment in place, but also by the cost of maintenance of the building and equipment, and protection from the elements.

9.2.3 Switchyard

To provide a reliable and flexible interface between the generating equipment and the power grid, a switchyard is usually associated with a hydroelectric powerplant. Switchyards include all equipment and conductors that carry current at transmission line voltages, including their insulators, supports, switching equipment, and protective devices. The system begins with the high-voltage terminals of the step-up transformer and extends to the point where transmission lines are attached to the switchyard structure. Switchyards are typically sited to be as close to the powerplant as space permits in order to minimize the length of control circuits and power feeders, and also to enable the use of service facilities in the powerhouse.

9.3 MAJOR MECHANICAL AND ELECTRICAL EQUIPMENT

Much of the major mechanical and electrical equipments installed in hydroelectric powerplants may be found in other generating, transmission, and distribution systems. Conventional types of power equipment are described in detail in other chapters of this handbook. In some cases, however, specialized equipment has been developed for hydropower applications. The following information is intended to emphasize equipment or configurations that are unique to hydropower facilities:

9.3.1 Turbines

The word “turbine” comes from Latin and means *spinning top*. Technically, hydraulic turbines that drive electric generators are called hydraulic prime movers. Whatever name is used, all hydraulic turbines convert fluid power into mechanical power by the same physical principle. They develop their mechanical power via the rate of change of angular momentum of the fluid. In most cases, the head is used to impart an angular momentum or prewhirl to the fluid. The action of the turbine runner is to remove this angular momentum or to straighten out the fluid streamlines. The effect of this change in angular momentum is to induce a torque on the shaft of the runner. The speed of rotation is the rate at which this angular momentum is changed, and torque multiplied by rotational speed is mechanical power.

The relative proportions of power transferred by a change of static pressure and by a change in velocity provide the most basic method of classifying turbines. The ratio of this transfer by means of a change in static pressure to the total change in the runner is called the degree of reaction, or more simply reaction. Therefore, if there is any significant pressure change in the runner of a turbine, it is a *reaction* hydraulic turbine. If there is no change in pressure, only in velocity, the degree of reaction is zero and these special cases are called *impulse* hydraulic turbines.

Aside from the most basic category as reaction or impulse, hydraulic turbines are classified in two separate ways—by the type of runner and by the configuration of the water passages. For reaction

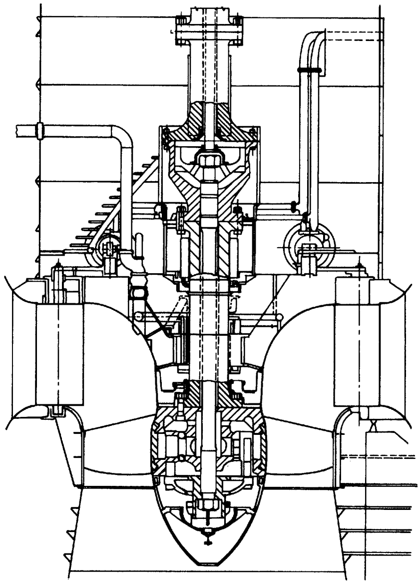


FIGURE 9-2 Sectional elevation of an adjustable-blade propeller (Kaplan) turbine.

turbines, there are different classifications of runners—axial, radial, and mixed. These terms denote whether the flow enters the runner parallel or perpendicular to the shaft, or at some angle in between. In modern reaction turbines, the flow leaves the runner axially.

For the lowest head applications, reaction turbines with propeller type runners are utilized. These may be *fixed blade* or if the pitch angle of the blades can be adjusted, they are called *Kaplans* (Fig. 9-2). In propeller turbines, the fluid enters and leaves the runner axially; therefore, these are axial flow machines. The ability to change the pitch angle maintains high efficiency over a wider power range. This is because as the flow rate is increased, or the head is increased, the velocity vector or the angle at which the fluid streamlines enter the runner gets steeper. Therefore, if the angle of leading edge of the blades is increased to remain aligned with the steepened fluid velocity vector, a higher efficiency is maintained. A cam in the governor that positions the blades based on the wicket gate opening controls the pitch angle of the blades. There are different cams for different increments of head. However, if instead of increments of head, the cam is also continuous in head; this is referred to as a *3-D cam*—the three dimensions being blade angle, wicket gate opening, and head.

A variation of the propeller design where the blades are not mounted perpendicular to the shaft, but at a downward or dihedral angle is the *diagonal* or *Deriaz* turbine. This arrangement transforms the runner into a mixed flow runner. The principle advantage in this arrangement is that it allows higher permissible operating heads.

Propeller, and especially Kaplan, turbines require a considerable amount of submergence under the tailwater elevation as they are prone to cavitation. In a Kaplan, maximum runaway speed occurs when the blades are full flat. (Full flat blade runaway speed can approach 300% of synchronous speed.) In order to minimize the runaway speed, the blades are normally hydraulically designed to drift to a full steep angle upon loss of governor oil pressure. However, maximum discharge at runaway speed is with the blades full steep (up to 150% of maximum discharge at synchronous speed).

A recent modification of the traditional Kaplan design is called a *minimum gap runner* (MGR). In this design, gaps between the blades and runner hub are hydraulically hidden and the discharge ring is a spherical cavity rather than a cylindrical cavity to minimize the gaps at the outer edge of the blades at steeper angles. The purpose of minimizing these gaps is to reduce injury to downstream migrating fish that will pass through the turbines.

For intermediate head applications, the most commonly used reaction turbine is the *Francis* turbine (Fig. 9-3). Depending on the exact shape of the inlet to the buckets, this may be a mixed or radial flow runner. A Francis runner looks somewhat like the impeller of a centrifugal pump. It has no adjustable or moveable parts. Unlike propeller or Kaplan turbines, where flow increases with runaway speed, Francis turbines tend to choke or reduce the flow with runaway speed. This characteristic can produce unwanted pressure rises in the penstock immediately following a load rejection (i.e., the loss of an electrical load).

For the highest head applications, the preferred choice is an impulse turbine. There are a number of different designs of impulse turbine runners. The most common is the *Pelton* (Fig. 9-4). In this design, jets discharge directly into buckets mounted around the periphery of a runner, which is housed in an atmospherically vented casing. Because the runner is at atmospheric pressure, impulse turbines are not subject to cavitation. The jet strikes a splitter in the middle of the bucket, which divides the jet in two. Each half of the jet turns almost a full 180° in the bucket and then falls free.

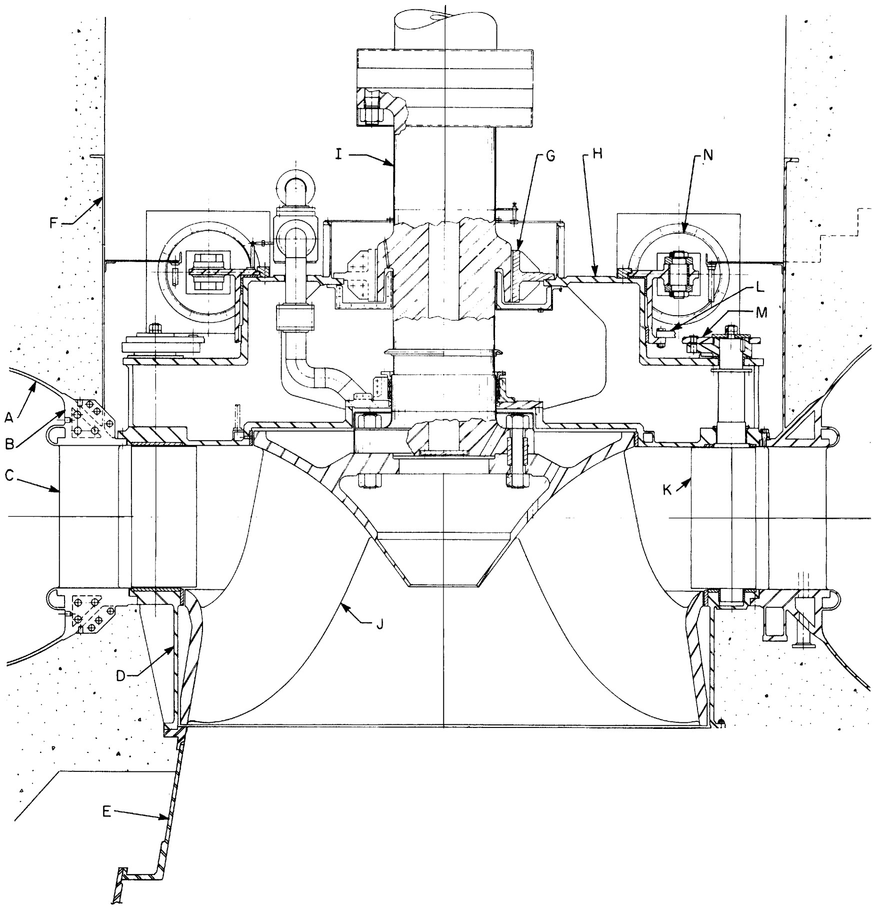


FIGURE 9-3 Sectional elevation of a Francis reaction turbine: A—spiral case; B—stay ring; C—stay vane; D—discharge ring; E—draft tube liner; G—main-shaft bearing; H—head cover; I—main shaft; J—runner; K—wicket gates; L—links; M—gate levers; N—servomotors.

The jet discharge is throttled or controlled by needle valves. Since this provides for a wide range of discharge from an individual nozzle and since multiple nozzles may be used on the same runner, Peltons can have a high efficiency over a very wide power range. If the shaft is mounted in the vertical, any practical number of nozzles can be used. However, if the shaft is horizontal, only two or three nozzles can be used. This is because of the need for gravity to clear the water from a bucket before the jet from the next nozzle strikes it.

A variation of the basic Pelton design is the *Turgo* impulse turbine. In this design, the jets strike the buckets at a side angle and discharge out the opposite side. The buckets do not have a splitter. The advantage is that this design allows larger nozzles with higher flow rates to be used for a given diameter of wheel.

Another design of impulse turbine is the *cross-flow* turbine. Today's cross-flow designs are developed from an earlier version called the *Banki* or *Michell* turbine. The name cross-flow comes from the action of the fluid to enter the vanes on one side of the horizontally mounted cylindrical runner and purported travel across the interior center and out the vanes on the other side. In point of fact, research has shown that the water actually rides around the periphery of the runner in the vanes until it can

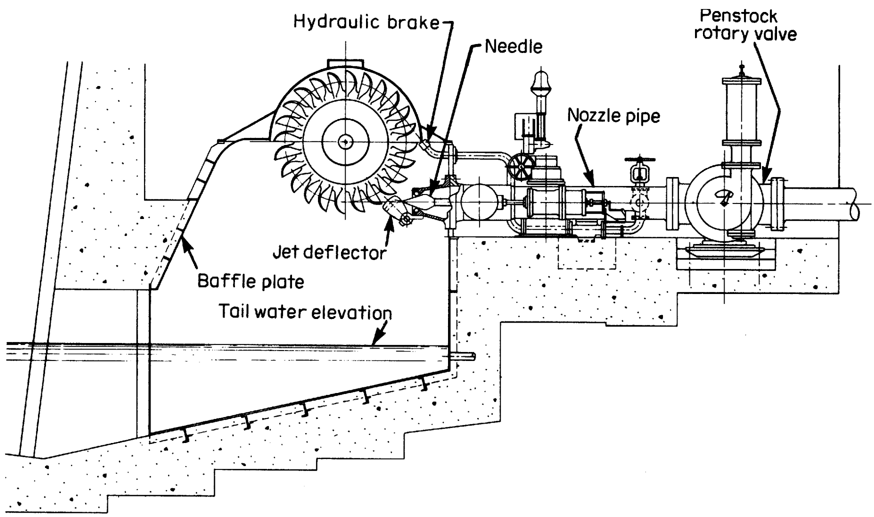


FIGURE 9-4 Section through a horizontal impulse turbine.

discharge out the other side. The principal advantages of this design are that it can operate at much lower heads than a Pelton and has a very wide range of flows. The wide flow range is achieved by dividing the runner into compartments. One commercial cross-flow turbine advertises a flow range of 16% to 100%. This is on the order of at least twice the flow range available from reaction turbines.

One significant difference between reaction and impulse turbines is that reaction turbines have draft tubes to convey the discharge from the runner to the tailrace. A draft tube is actually a conical diffuser, in which the cross-sectional area continually expands with distance along the centerline. The purpose of a draft tube is twofold. The first is to confine the high velocity discharge under the runner so that the static pressure may be below atmospheric. This increases the head across the runner. The second is to slow that high velocity prior to discharge into the tailrace. As a consequence of slowing the velocity, the pressure is recovered. For this latter reason, draft tubes are sometimes referred to as pressure recovery devices.

Aside from the different types of runners, turbines are classified by the different configurations of their water passages. Reaction turbines typically have vertical shafts. The runners of propeller type turbines with vertical shafts are surrounded by a circular water passage called a semispiral case. This is generally formed by concrete and fed with water directly from the forebay through intake bays. Francis turbine runners are surrounded by a full spiral case and, because of the higher head and increased water pressure, this is generally formed from rolled steel plate and then embedded in concrete. Water is generally conveyed to these spiral cases through penstocks. Typically, just upstream of the turbine there is a shut-off or isolation valve in the penstock. When this valve is closed, the turbine can be dewatered. Spiral cases supply water to circular sets of wicket gates and stay vanes in what is called the distributor section. The wicket gates control the rate of flow. The principal purpose of the stay vanes, however, is structural rather than hydraulic. They are used to transfer the vertical load of the weight of the upper powerhouse structure to the powerhouse foundation. Stay vane design may improve the efficiency of the turbine by providing smooth transition of flow to the turbine runner. With a vertical shaft, the beginning of the draft tube under the runner is pointed downward. In order to minimize the amount of required excavation, draft tubes are often constructed with an elbow to turn them horizontal about mid length and these are called elbow draft tubes.

To reduce excavation and cofferdam costs, low head units may have horizontal or inclined shafts. The water passages for horizontal or inclined shafts have less severe bends and turns and, therefore, tend to have lower hydraulic losses and higher efficiency. A common horizontal shaft configuration is to house the generator upstream of the runner in a submarine-like bulb. These are called *bulb* turbines, even though the runners are usually conventional fixed blade propeller or Kaplan types (Fig. 9-5). A variation on this design is to house the upstream generator in a concrete silo with the water passages on either side. This is called a *pit* turbine. Pit turbines typically use speed-increasing gearboxes to reduce the size of the generator. Rather than the generator being upstream, the shaft may extend downstream, either horizontally or inclined at an upward angle. In these configurations, the shaft can extend through the draft tube liner so that the generator is not housed inside the water passages. Whether the shaft is horizontal or inclined, these are referred to as *tubular* turbines (Fig. 9-6). There is even a design where the generator is housed around the periphery of the runner, called a *rim* turbine.

Due to the higher head, water is conveyed to impulse turbines through penstocks. The runners of most impulse turbines rotate in some type of splash containing housing. Since the runners of impulse turbines are vented and operate at atmospheric pressure, they must be set at an elevation higher than the maximum tailwater elevation to avoid being flooded out. The discharge is conveyed to the tail-race through some type of open surface canal or tunnel.

9.3.2 Generators

A hydraulic turbine converts the energy of flowing water into mechanical energy; a hydroelectric generator converts this mechanical energy into electricity. Almost all hydroelectric generators are synchronous alternating-current machines with stationary armatures and salient-pole rotating field structures. The stationary armature (stator) is comprised of a steel core encircled by a frame that is mounted to the powerplant foundation. A 3-phase armature winding, in which the alternating current is generated, is embedded in the stator core. The three phases of the armature winding are Y-connected at the neutral end. The rotating magnetic field is typically produced via a direct current–excited winding connected to an external excitation source through slip rings and brushes. An amortisseur winding is often mounted on the rotor poles to dampen out mechanical oscillations that may occur during abnormal conditions. The stators of hydroelectric generators usually have a large diameter armature compared to other types of generators, and can exceed 60 ft. The capacity of hydroelectric generators may range from a fraction of an MVA to more than 800 MVA. Hydroelectric generators are typically air-cooled, although the stator windings of the highest-capacity machines may be directly water-cooled.

The electrical and mechanical design of each hydroelectric generator must conform to the electrical requirements of the power transmission and distribution system to which it will be connected and also to the hydraulic requirements if its specific plant. Such constraints have made it impossible to standardize the size or capacity of hydroelectric generators. The rotational speeds of the generator and turbine are usually the same because their shafts are directly connected. In some cases, however, a speed increaser (gearbox) is used to enable the generator to operate at a higher speed than that of the turbine, thus permitting a smaller and less expensive generator to be used. Hydrogenerators are relatively low-speed machines, typically ranging from 50 to 600 revolutions per minute (rpm). Large diameter units with a lower hydraulic head operate at slower speeds, whereas physically smaller units with high hydraulic head operate at higher speeds. The best speed for each type of turbine is first established, and a generator is then designed that will produce 60 cycle alternating current at that speed. For a generator operating in a 60-Hz system, the rotational speed (in rpm) times the number of field poles on the rotor is always 7200. Hydroelectric generators are normally vertical shaft machines, although some smaller units are mounted horizontally.

9.3.3 Governors

Almost all hydraulic turbine generator units run at a constant speed. The governor keeps each unit operating at its proper speed through a high pressure hydraulic system that operates wicket gates which control water flow into the turbine. When there are load changes or disturbances in the power grid, the governors respond by increasing or decreasing power output of the generating units to meet power demands and keep the frequency of the power grid at 60 cycles. Governor-operating characteristics will be determined from the electrical, mechanical, and hydraulic characteristics of

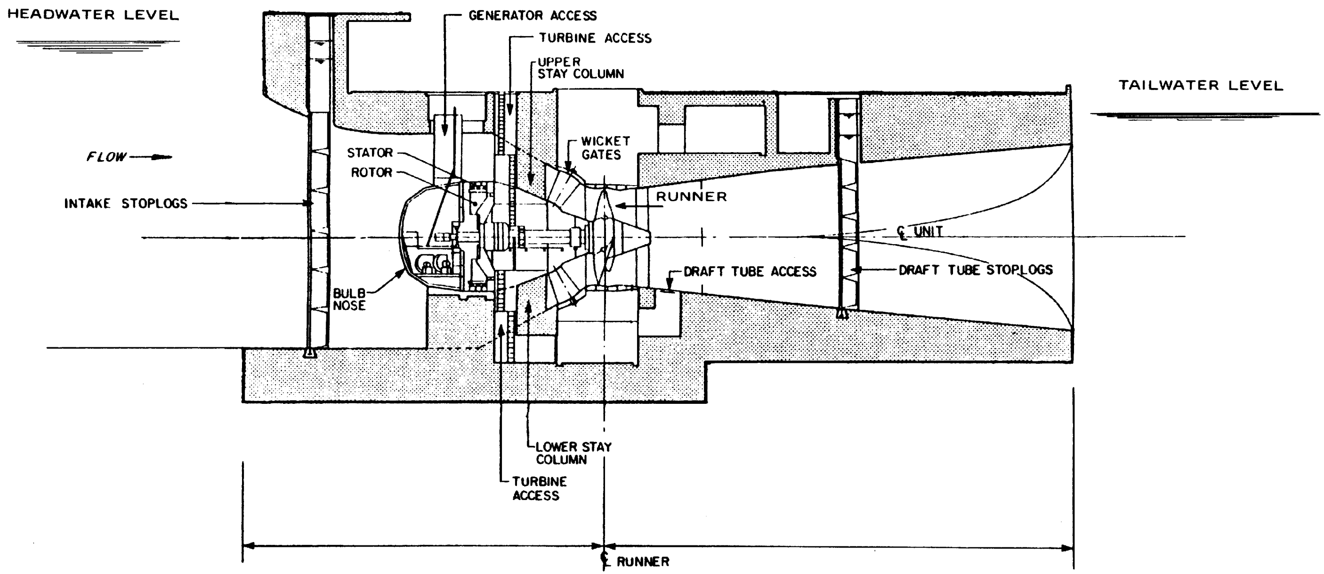


FIGURE 9-5 Sectional elevation of an axial-flow (bulb) turbine.

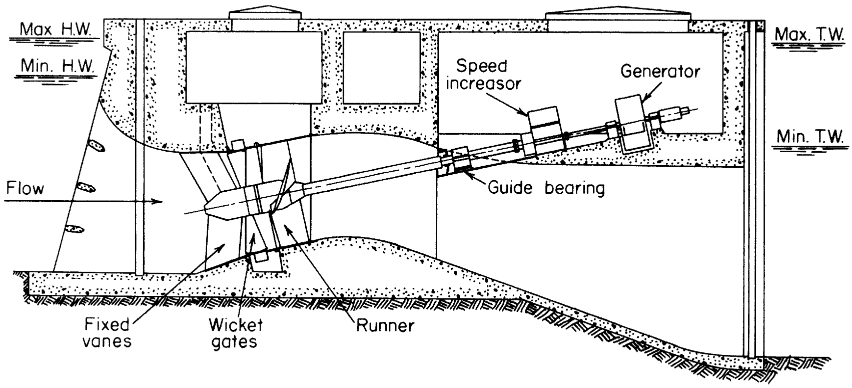


FIGURE 9-6 Sectional elevation of an axial-flow (tubulas) turbine.

the generator, turbine, and penstock. Older governors use mechanical speed sensing and control, interfaced to the hydraulic system to govern turbine speed. Newer systems incorporate electronic or digital speed sensing and controls with a hydraulic interface to the turbine governor.

9.3.4 Excitation Systems

The function of the excitation system is to supply direct current to the field winding of the main generator. This current is used to create the rotating magnetic field necessary for generator action. Control of the current in the field winding must be accurate, sensitive, and reliable to allow stable and economic operation of the generator. All excitation systems include an exciter, a voltage regulator, generator voltage and current transformers, and limiters and protective circuits. The exciter may be a rotating type that is directly connected to the generator shaft or a modern static system utilizing solid-state devices fed from a high-voltage bus.

9.3.5 Circuit Breakers

A circuit breaker is a mechanical switching device, capable of making, carrying, and interrupting current during normal operating conditions as well as under specified abnormal conditions, such as during a short circuit. Circuit breaker ratings and location are considered during the preliminary design of a powerplant to meet the switching flexibility and protection requirements of the generators, transformers, buses, transmission lines, etc. Generators at large multi-unit powerplants are commonly configured so that a dedicated unit breaker is situated between the phase terminals of each generator and the main step-up transformer. Smaller plants may only have provision for switching via a switchyard breaker on the high voltage side of the step-up transformer, the generator and transformer being connected and disconnected to the transmission network as a unit. In some cases, circuit breakers are used to perform switching between a main and transfer bus in the switchyard. A variety of switching schemes are possible and commonly used, depending on the local requirements and economic considerations. The ratings, design, construction, and operation of circuit breakers installed at hydroelectric powerplants are generally similar to those used in other power system applications.

9.3.6 Transformers

Most dams and associated hydroelectric powerplants are located a great distance from population centers; therefore, the economics of transmitting power over long transmission lines must be considered. Traditionally, hydro generation has been in the medium voltage range, or about 15 kV. Power transformers step the voltage up to the 100 to 500 kV range for a more economical transmission from the powerplant by minimizing transmission line losses.

Transformers associated with hydroelectric generation may differ somewhat from those used in transmission and distribution applications. For example, it is not uncommon for a single step-up transformer to accommodate multiple hydro generators. To maintain fault isolation between generators for such a transformer-sharing arrangement, each machine may be connected to an exclusive primary winding. Multiple primary windings are often used in hydropowerplants because of the relatively small power output ratings (MVA) of a typical generating unit. Thus, a single large transformer can be sized and manufactured to meet the requirements of multiple generators, providing a substantial savings in equipment cost.

Also unique to hydro plants is the use of the forced-oil-water (FOW) transformer cooling method. Although few, if any, new transformers are cooled this way because of environmental issues, the availability and efficiency of FOW made it the method of choice in the past. The availability and proximity to water made FOW an attractive and unique solution to step-up power transformer cooling.

9.4 BALANCE OF PLANT

9.4.1 Station Service

The station service supply and distribution system is provided to furnish power for the plant, dam auxiliaries, lighting, and other adjacent features of the project. Since hydroelectric plants are capable of starting with relatively low auxiliary power needs (compared to steam plants), they are often used to provide "black start" capability for the local transmission system. If the plant is to provide this capability, the station service system design must include an automatic start engine-driven generator to provide power to critical auxiliary powerhouse loads. This is in addition to the engine-generator the plant must have to operate spillway gates and other river regulating works when offsite power is unavailable.

The complexity and operational flexibility of the station service system are related to the number of main generator units and the importance of the plant to the overall power system. Large plants with numerous units may have two station service transformers and even dedicated station service hydro generators. Station service transformers are often fed from different main generator unit buses to allow the main units to carry station service loads upon disconnecting from the system. Smaller hydroelectric plants may have only one station service transformer and an engine-driven generator.

9.4.2 Switchgear

Station service systems at hydroelectric plants utilize standard metal-clad switchgear assemblies. In large plants, where the distance between the station service switchgear and the utilization equipment is large, the use of 4.16 or 13.8 kV distribution circuits is used where economically justified. Double-ended switchgear, consisting of two dry-type 13.8 or 4.16 kV transformers fed from separate sources, and connected to 600 V switchgear with a normally open tiebreaker between the two sections, is often used for important load centers.

9.4.3 Controls

Plant controls are comprised of computer-based controls, hard-wired or programmable logic, indicating and recording instruments, protective relays and similar equipments. Each generator or pair of generators often has local control panels or switchboards located near the units. For multiple unit plants, centralized controls are also used to coordinate the operation of all units within the powerhouse. The centralized control equipment is situated in a control room located at an elevation above the maximum high water levels. Centralized control is used to apportion MW and MVAR loading among multiple machines while respecting machine operating limits.

Small plants may not have a dedicated control room. They may have the local unit control panels integrated with the station service switchgear lineup, which usually requires additional compartments to accommodate the needed equipment.

9.4.4 Instrumentation

The instrumentation at hydroelectric projects has a number of unique features, most of which involve the measurement of hydraulic and mechanical parameters. There are two basic types of these parameters—performance and positional. Performance parameters include power, head, and flow. Positional parameters refer to such items as wicket gate opening, nozzle jet opening, and Kaplan blade angle position. Instrumentation to measure generator power output is covered in other chapters of this handbook.

Head is a performance parameter that can usually be measured to a high degree of accuracy. The traditional method is by the use of stilling wells. These are vertical tubes with restricted openings under the water surface of the elevation to be measured. This restricted opening serves to dampen the effect of wave action and provides a steady water surface inside the well. Floats with counterweights or electro tapes can be used inside the stilling well to measure the water surface elevation. More modern measuring devices use radar or acoustic waves rather than stilling wells. These waves are bounced off an open water surface to measure a vertical distance between the instrument and water surface.

The head measurement described above presents some challenges. Although head refers to distance or height, it is used to express the pressure resulting from the weight of a body of liquid since the weight is directly proportional to the height. Therefore, head actually represents a difference in hydraulic energy levels. However, when water is flowing, the elevation of the water surface is not the true energy level because it does not account for the kinetic energy contained in the velocity head, $V^2/2g$. Thus, measuring the elevation of a tailwater surface at a draft tube exit does not provide a correct downstream energy level. In addition, bends in the river or the operation of adjacent units may cause the head on any one individual unit to differ from the location where head is measured for the powerhouse. Thus, the location where head is measured is a unique feature of the accuracy of head measurement.

Volumetric flow rate is generally the most difficult performance parameter to measure to any degree of accuracy. For projects with penstocks or at least a water passage with a constant cross section of sufficient length, there are several methods to accurately measure absolute flow. However, with large, run-of-the-river projects where the cross section of the water passage is continually changing, accurately measuring flow becomes very difficult. In such situations, relative flow may be measured instead to determine a relative efficiency. Relative flow is uniquely measured by determining the effect that absolute flow has on another parameter that can be measured. The Winter-Kennedy piezometer system is commonly used for this purpose. This consists of two piezometer taps, one on the inside and the other on the outside of the spiral or semi-spiral case. The square root of the piezometric or pressure difference between these two taps is directly proportional to the flow rate. Therefore, a relative efficiency of the turbine-generator may be measured as $E_{rel} = \text{kW}/(H\sqrt{D})$, where kW is the generator output in kilowatts, H is head in feet, and D is the piezometric difference, usually in feet.

With reference to the positional parameters, the actual wicket gate opening is defined as the dimension of the largest sphere that can pass between the two gates. When a turbine is unwatered, the gate opening may be calibrated with a curved scale on the wicket gate operating ring or even an angle indicator on the top of the wicket gate stems. However, in order to use a straight-line motion sensor to measure the amount of wicket gate opening, the stroke of the wicket gate servomotor reach rod is used. This measurement is often called gate opening and used directly without converting to actual gate opening, even though the two do not have a linear relation. Because of the curved shape of wicket gates, the relation between actual gate opening and servomotor stroke is a shallow "S" curve. In addition, at each end of the servomotor stroke there is an area of squeeze. This is where the reach rod is moving to take up slack in the linkages, but the gates are in contact or at their stops and not moving. Therefore, a reading of a gate opening tends to be unique to each project.

The inner oil pipe in the oil head on a Kaplan turbine is generally used as the indicating surface to measure blade angle. This provides a linear motion for a position sensor and can be calibrated from the master blade position ring on the hub when the unit is unwatered. However, each turbine manufacturer has a different trigonometric convention to define the actual blade angle. There is no industry standard or convention for this measurement. Therefore, a reading of a blade angle tends to be unique to each family of turbines in a powerhouse.

9.4.5 Protection

Hydroelectric plants are protected using standard generator, transformer, and distribution system protection methods and schemes. A few features or practices that may not be common to other types of plants are discussed here. On large generators, whose stator windings consist of multiple-turn coils with multiple parallel circuits per phase, split-phase differential relaying is sometimes used to provide increased sensitivity to turn-to-turn shorts. Under-frequency and over-frequency relaying is often not used, or is set very liberally compared to steam units as the hydraulic turbine and generator are not susceptible to damage due to off-nominal frequency operation. Special schemes are used to provide selectivity on isolated-phase bus ground faults in installations where multiple high-resistance grounded units are tied together at the generator terminal voltage level.

9.4.6 Direct Current Systems

A direct current (dc) system is used to provide independent power for auxiliary equipment and systems including controls, relaying, data acquisition, communication equipment, fire protection, inverter, generator exciter field flashing, alarm functions, and emergency lights. The DC system consists of a storage battery with its associated charger, and provides the stored energy required to ensure adequate and uninterruptible power for critical powerplant equipment. In the event of a complete loss of station service power, the dc system supplies the power needed to conduct an orderly shut down of generating equipment which could be damaged if operated without auxiliary systems such as control power, cooling water, lubrication oil, etc. An inverter is fed from the battery for the critical alternating current loads. For plants equipped with black start capability (i.e., the ability to start up a plant when separated from the transmission system and the generators have been shut down), the dc system provides a dedicated power source for auxiliary equipment as well as field flashing of the generator exciter in order to restore a small amount of residual magnetism in the generator exciter field to allow the generator to build up voltage during start-up.

9.4.7 Annunciation

Annunciation systems are used to alert someone (typically the control room operator) when a critical plant or equipment parameter falls out of tolerance and requires attention and/or action. Annunciators generally provide visual and audible signals, such as lights and flashers along with a horn, bell, or buzzer. Acknowledge and reset functions may also be provided. Annunciation systems may consist of a separate annunciator hardwired into the plant, or a software feature programmed into the central control system. Typical alarm points include turbine bearing oil trouble, unit bearing overheating, generator excitation system trip or trouble, generator cooling water flow, generator stator high temperature, governor oil trouble, transformer differential, and transformer overheating.

9.4.8 Miscellaneous Equipment and Systems

A wide variety of mechanical and electrical auxiliary equipment and systems may be found at hydroelectric powerplants. Of the following items, not all will be incorporated into all plants. The size, service, and general requirements of the facility will usually determine which items are needed: water supply systems for raw, treated, and cooling water; sewage disposal equipment; heating, ventilating, and air-conditioning systems; fire detection and protection systems; telephone and code call systems; elevators; intake and discharge gates and valves; penstock drainage and unwatering systems; station drainage system; air receivers for draft tube water depressing system; insulating and lubricating oil transfer, storage, and purification systems; compressed air systems for service, generator brakes, and turbine governor; emergency engine-driven generators; metal-enclosed buses, surge protection equipment; and transformer oil pumps.

9.5 DESIGN ASPECTS

9.5.1 Criteria and Philosophy

The basic approach to designing a hydroelectric project is to first determine the rated discharge of the powerhouse. Hydrologic or other records are used to develop a historical graph of the frequency of volumetric flow rates. The flow values may be mean daily, weekly, or monthly. The period of record should be as long as possible. From this information, a flow exceedence graph is developed. This is a graph of flow versus the percent of time that a flow value exceeds. As a general rule of thumb, run-of-the-river projects (those having little reservoir storage capability) are sized to a Q_{20} and projects with storage are sized for a Q_{30}. The term Q_{20} means a flow value that exceeds 20% of the time. In other words, the project could utilize all of the flow for generation 80% of the time. Similarly, Q_{30} means a flow value that exceeds 30% of the time.

Next, a design head is determined. This is different from rated head and is the head at which best efficiency is to occur. Such a determination depends on the specifics of the project, but a weighted average head is often used. With the hydraulic head and estimated hydraulic losses in the penstock, a power duration curve may be developed. Annual energy production may then be calculated from the area under this curve multiplied by an appropriate conversion factor.

With the value of design head, a dimensionless parameter known as specific speed is determined from a historical experience curve of specific speed versus design head. Specific speed is defined as the speed at which a turbine would rotate if it were 1 ft in diameter and operating under 1 ft of head. It is calculated in U.S. units as $N_s = N(\mathrm{HP})^{1/2}/H_d^{5/4}$, where N_s is the specific speed, N is the rotational speed in rpm, HP is the turbine output (at full gate in this instance) in horsepower, and H_d is design head in feet. The specific speed is used as a classifying parameter of hydraulic turbines. With the value of specific speed, the type and configuration of turbine can be determined, which historically has been found to be the best selection for the same conditions.

Next, the size and number of generating units required to pass the rated discharge is determined. Generally, a fewer number of larger units is the more economical option. For some projects, the physical size of the unit has been limited to the maximum size runner that could be shipped in one piece. This is largely due to the extra manufacturing costs involved in furnishing split runners. However, other considerations, such as flexibility of operation and minimum loss of capacity during shutdown for repair or maintenance, may dictate the use of more, smaller units. Sometimes to achieve increased flexibility with few units, different size units are used in the same powerhouse.

9.5.2 Ratings

In the design of a hydroelectric plant, the generating equipment is first sized and then afterwards it is rated. Sizing refers to selecting the physical size of the equipment. Generally, hydrologic considerations of head and flow provide the basis for the determination of the type, number of units, runner diameter, setting, and synchronous speed of each hydraulic turbine selected for a particular project. Then the generator is sized to match the turbine speed and expected output at a selected head.

Once the equipment is sized, it can then be rated. However, the rating is done in the reverse order. First, the purchaser usually specifies the temperature rise criteria from which the manufacturer then rates the generator. The generator is rated in terms of kVA and power factor. It is a standard practice to set such thermal limits to that power output which causes no more than a 60 or 80°C temperature rise above ambient in the generator windings. It is a "soft" limit in that it can be physically exceeded. However, this can have a detrimental effect on the operational life of the insulation. Converting this generator output rating into a generator horsepower input gives the rating of the turbine. However, the actual turbine rating is not in units of horsepower, but in units of feet of head. This is because as head is increased a turbine can produce more power. This head is usually the net head, or the same head as used to develop the turbine performance characteristics. Therefore, the rating of a turbine is actually the lowest net head at which the turbine can drive the generator to produce its rated electrical output. This is a unique point, with the wicket gates full open. Therefore, it is also a "hard" limit since the turbine cannot physically produce more power unless the head is increased. As a consequence, if the turbine and generator are procured separately, it is not the turbine manufacturer but the purchaser who actually establishes the turbine rating. The discharge at this rated head is referred to as rated discharge.

There are a couple of notable alternatives to this rating procedure. Some owners equate the turbine rating to the generator nameplate rating at unity power factor, regardless of the nameplate power factor. This is referred to as the generator *capacity*.

Seeking to reduce the cost of procuring hydroelectric generators, a number of years ago some purchasers began rating their turbines at 115% of the generator nameplate rating at nameplate power factor. They then specified that the generators must be capable of "continuous operation" at 115% of rated generator output. In other words, they rated a turbine at the generator's overload rating and then specified the generator overload rating the way the regular nameplate rating is usually specified. (The turbines were also required to be able to produce an output of 115% of generator nameplate at unity power factor, at a higher head, without exceeding mechanical limits.) Today, the standard procedure is to use the full overload rating as the nameplate rating.

9.5.3 Speed Settings

Although some extremely small hydraulic turbines may power induction generators, most turbines are directly connected to synchronous ac generators. As a consequence, the speed of rotation must agree with one of the synchronous speeds required for the system frequency. The prevailing frequency for most systems in the United States is 60 Hz. In Europe and certain other parts of the world, 50 Hz is used. Synchronous speeds are determined by the formula, $N = 120 \times$ frequency (in cycles per second)/number of poles in the generator. The number of poles must be an even number since the poles are in pairs.

The need to rotate at a synchronous speed means that the turbine is constrained to rotate at a single speed as the hydraulic conditions of head and stream flow vary. This is a major constraint unique to the design of hydropower, negatively affecting the efficiency and smooth operation of the turbine. The speed should be as high as practicable since this decreases the cost of the turbine and generator. The proper selection of the synchronous speed is usually done with reference to the specific speed of the hydraulic turbine. As defined in Sec. 9.5.1, specific speed is calculated as $N_s = N(\mathrm{HP})^{1/2}/H_d^{5/4}$. However, in this calculation, the turbine output value is the horsepower at peak efficiency at design head rather that at full gate at design head. With the values of turbine output and design head, and the turbine specific speed at peak efficiency, a trial rotational speed is calculated. This is usually rounded up to the next higher synchronous speed, and the turbine output at peak efficiency recalculated. If the turbine output at either peak efficiency or full gate or rated discharge is not as desired, the size of the turbine is changed and the speed calculation repeated.

9.5.4 Water Hammer and Mass Oscillations

Water hammer is a transient pressure phenomenon that can occur in moving water in a closed conduit. If there is a change of velocity, for example, due to the closing of a valve, pressure waves are created that travel up and down the conduit. Upstream of the closing valve, the fluid is progressively decelerated and compressed, causing a positive pressure transient to travel upstream. Once this wave reaches an open surface, it is reflected back toward the valve as a negative pressure wave. When it reaches the valve, it will be reflected again. This process is repeated time and again until the wave is attenuated by friction. On the downstream side of the valve, the transients are reversed with a negative wave initially traveling downstream. The greater the distance between the valve and an open surface, the higher the peak magnitude of the pressure rise. The pressure waves travel at the speed of sound or the acoustic velocity called the celerity, which is given the symbol "a." The celerity varies depending on the conduit boundaries, the static pressure, and the water temperature (and salinity), but is typically on the order of 2000 to 3000 ft/s.

The water passages of a hydroelectric project must be designed to withstand both the maximum positive and negative pressure transients to prevent potentially catastrophic damage to the valves or rupture of the penstock. The magnitude of the maximum transients can be controlled by the design of the dynamic elements. For example, the maximum rate of valve movement, such as the rate at which the wicket gates of a turbine can move, can be controlled by limiting the size of the oil ports in the servomotors. Slower gate closure, however, results in higher generator overspeed when an electrical load is lost.

The magnitude of pressure transients can also be mitigated by surge tanks, accumulators, or quick acting pressure relief valves. A surge tank has an open surface. Consequently, it provides a partial negative return wave and acts to shorten the effective length *L* of the conduit. There are a number of different types of surge tanks, such as the simple riser, restricted riser, differential, etc. The more the flow is restricted in either or both entering and leaving the surge tank, the less the pressure transients are mitigated, but the more hydraulically stable the surge tank. An accumulator does not have an open surface, but has an enclosed dome of air or gas. The quick acting valves operate in a manner analogous to safety valves.

A separate but related phenomenon to water hammer is mass oscillation. Rather than a wave within the fluid, this term denotes the actual movement or velocity of the fluid. Consequently, it is much slower and, therefore, separated from pressure transients on a time basis. Using a surge tank as an example, if the turbine wicket gates start to close, a positive pressure wave is transmitted upstream. Part of that wave continues upstream to the reservoir water surface, but a part also reaches the free surface of the surge tank and is reflected back as a negative wave. As the initial positive wave reaches the riser of the surge tank, it causes part of the flow still coming downstream to be diverted up into the surge tank and the elevation of the water surface in the tank starts to rise. Consequently, the deceleration, $-dV/dt$, of the flow upstream of the tank is not as rapid, which serves to mitigate the magnitude of the pressure transient upstream from that point. Another example of mass oscillation is the separation of the water column. If a valve is closed very rapidly in a high velocity flow, the momentum of the fluid downstream of the valve can cause the fluid column to separate at that point. As can be imagined, this can have dire consequences.

9.6 OPERATIONAL CONSIDERATIONS

9.6.1 Runaway Speed

Runaway speed is the maximum rotational speed to which a generating unit can be driven with an open circuit breaker and the available hydraulic and mechanical conditions. The term usually refers to a fixed wicket gate opening, and in the case of Kaplan turbines, a fixed blade angle. During a load rejection, the water column continues to provide energy to the turbine runner. Since this energy can no longer be converted into electrical energy, a portion is mechanically stored and the rest is dissipated in turbulence before being discharged from the turbine. The energy is stored via increased angular momentum of the turbine runner, shaft, and generator rotor. The total amount of energy that can be stored is a function of the rotating inertia, or WK^2, and the increase in rotational speed.

If the wicket gates do not move to a closed position, the speed will increase until limited by hydraulic conditions, windage, and friction. The hydraulic conditions include the available head, the turbine's performance characteristics such as off design efficiency, and cavitation, which can reduce the efficiency of the energy transfer. Windage refers to air resistance, mostly in the generator, and friction refers to mechanical "sliding" friction. Ultimately, the decreased amount of fluid energy that can be transferred from the water column is balanced by the increased windage and friction, at which point runaway speed is achieved. The higher the head, the larger the wicket gate opening, or the flatter the blades on a Kaplan turbine, the higher the runaway speed. For this latter reason, the blades on Kaplan turbines are often designed to tilt to their steepest position on loss of governor control.

Francis and Kaplan turbines have different runaway speed characteristics. Francis turbines typically have less WK^2 than Kaplans and, therefore, achieve runaway speed faster. At runaway speed, Francis turbines tend to "choke" the flow, reducing the discharge. Kaplans, on the other hand, tend to increase the flow with increasing speed at a given gate and blade angle. On Kaplan turbines, on-cam runaway speed is achieved if load is rejected, the gates do not move, and the blades are at the proper cam position for the gate opening and head and do not move. If the blades move to any other position without moving the gates, it is referred to as off-cam runaway speed. Also with Kaplans, maximum runaway discharge is when the blades are full steep.

If the wicket gates are free to move and the unit is under governor control, a shutdown sequence is initiated upon load rejection. However, since the wicket gates take a finite time to close, a transient increase in synchronous speed, known as overspeed, is achieved. In order to limit this overspeed, the wicket gates should close as quickly as possible. However, the faster they close, the higher the pressure transient of water hammer that is sent back upstream. For this reason the rate of closure, called the gate-timing element, is a compromise between overspeed and water hammer. The peak overspeed can be reduced by increasing the inertia of the rotating parts. Normally, the maximum design overspeed is 150% of synchronous speed.

9.6.2 Cavitation

Cavitation is a phenomenon involving the creation of bubbles containing water vapor. It occurs when the local pressure is reduced to or below the vapor pressure of water. Literally, the water boils, but at low temperature. The formation of vapor-filled bubbles is more likely to occur under conditions of high flow velocity (such as high rpm operation or high flow rates) and low pressure (such as low tailwater). Cavitation occurs in reaction hydraulic turbines, but not in impulse turbines whose runners are vented to atmospheric pressure. Minimum pressures in reaction turbines tend to occur at the trailing edge on the underside or suction side of blades or buckets. As these bubbles are carried downstream, back to higher-pressure areas, they collapse or implode. These implosions generate extremely high pressure pulses, sufficient to pit and erode the surfaces of the hardest steels. There are many types of cavitation including leading edge, areal, traveling, leakage, etc. Cavitation damage reduces turbine-operating efficiency and, if left unchecked, can lead to severe damage and extensive repairs. Most types of cavitation, but not all, can be lessened or eliminated by increasing the submergence of the turbine runner.

Model tests are primarily used to check turbine runner, wicket gate, draft tube, casing, and sometimes inlet work designs for optimum performance. They are also used to predict the conditions under which cavitation will occur. However, their predications have a degree of uncertainty because cavitation is actually 2-phase flow and the same hydraulic model cannot have similitude with both a liquid and a gas phase. The result is that a model prediction of cavitation is usually biased. That is, if the model shows cavitation at a certain condition, the prototype will definitely cavitate at that same condition. However, if the model does not cavitate, the prototype may still experience cavitation. For this reason, turbine designers try to maximize the amount of safety margin.

Aside from submergence, controlling cavitation is best achieved through design of the runner so that velocities at critical areas do not lower the static pressure to the vapor pressure. Other control methods include welding an overlay of a cavitation-resistant material on the base metal. Sometimes, special anticavitation fins are added to turbine blades on propeller type turbines to minimize blade tip cavitation. The injection or aspiration of air bubbles has been used to cushion the action of the pressure pulses.

9.6.3 Turbine Efficiency

The formula for turbine efficiency is developed from the definition of fluid power. If the volumetric flow rate Q, in cubic feet per second (ft^3/s), is multiplied by the specific weight of water γ, in pounds (force) per cubic foot (lbs_f/ft^3), the weight flow rate γQ, in lbs_f/s, passing through the turbine is obtained. This term may then be multiplied by the head H, in feet. (Technically, head is called the specific energy E and has units of ft-lbs_f/lbs_f.) The resulting expression, γQH, has units of ft-lbs_f/s and represents the power available in the fluid column. Dividing this expression by 550 ft-lbs_f/s per horsepower gives the power in the fluid column in units of horsepower. Since this expression represents power "in," dividing it into the horsepower "out" of the turbine yields the turbine efficiency

$$E = HP/(\gamma QH/550)$$

If the combined turbine-generator efficiency is to be calculated, the formula may be changed to

$$E_{t-g} = 1.3411(\text{kW})/(\gamma QH/550)$$

where kW is the generator output in kilowatts.

In selecting the type of turbine for a given hydroelectric powerplant, it is important to consider the efficiency performance of the various types of turbines available for the head contemplated. Not only is this true for the value of the maximum efficiency obtainable, but also for both the percentage of full load where this maximum efficiency occurs and the efficiencies at part loads and at full load (Fig. 9-7).

Impulse turbines are usually a couple of percent less efficient than comparable reaction turbines. However, because of their ability to use multiple jets, they can have a flat efficiency profile over a very wide power range. Francis turbines can have among the highest peak efficiencies. However, their runners have no mechanical adjustment and, therefore, their profiles are sharply peaked with efficiencies degrading significantly at part loads. Fixed blade propeller turbines also can have high peak efficiencies, but like Francis turbines, their profiles are sharply peaked. This latter feature is modified by Kaplan turbines, which are propeller turbines with adjustable blades. As head and flow conditions change, the pitch angle of the blades can be adjusted to maintain a relatively flat efficiency profile over a wide range of power and head. However, Kaplan turbines do have slightly reduced peak efficiencies due to increased leakage around the ends of the blades.

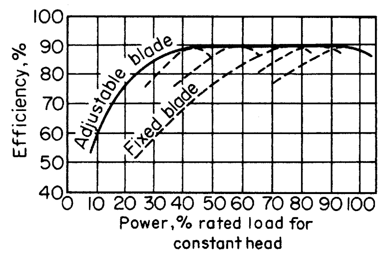


FIGURE 9-7 Efficiency-load relations for fixed- and adjustable-blade propeller turbines.

9.6.4 Operating Limits

Hydroelectric projects often operate under a number of different limits or constraints. These limits may affect either generating capacity or generating efficiency, and usually originate from one of three sources—physical, contractual, or regulatory.

Physical limits are those imposed by the physical characteristics of the generating equipment or the hydraulics. For example, the maximum output of the generator may be limited such that the temperature rise above ambient within the generator insulation does not exceed a specified value. The output of the turbine may be limited to avoid operating in zones where draft tube surging occurs. Such surging causes a fluctuation in power output. The head may be limited to a minimum value to prevent the forebay from being low enough to allow air to be drawn into the penstock through a vortex at the intake.

Contractual limits are imposed by the procurement specifications of the equipment. They usually apply while the equipment is under the manufacturer's warranty. For example, the specifications may limit the turbine output as a function of head to avoid cavitation damage while the turbine is under warranty.

The majority of operating limits are regulatory in nature. The hydropower licensing procedure described in the following section provides ample opportunities for the imposition of operating limits. These limits are of two types—static or time varying. An example of a static limit is a project that is in the path of ocean bound juvenile anadromous fish which may be restricted to operating within 1-percent of peak turbine efficiency during migratory seasons. (Water turbulence is at a minimum at peak turbine efficiency and, therefore, fish survival is thought to be increased.) An example of a time varying limitation is a limit on the ramp rate or the rate at which the generated power level may be changed. Such a limit may be imposed to prevent varying the elevation of the tailwater too rapidly.

9.6.5 Regulatory Requirements

Hydropower is regulated through three legal venues—water rights, state regulatory permits, and federal licensing. Water rights are required on all hydropower developments in the United States. These are administered through state statutes, which vary greatly from state to state. In addition, individual states have various other legal requirements, involving consultations with state agencies and permits.

Federal regulation of hydroelectric power began in 1920 when Congress enacted the Federal Water Power Act and established the Federal Power Commission (FPC) to administer the Act. In 1977, as part of the Department of Energy Organization Act, Congress created the Federal Energy Regulatory commission (FERC), which assumed most of the FPC's hydro regulatory responsibilities. This commission has jurisdiction over nonfederal development of hydropower projects, constructed after 1920, which meet one or more of the following criteria:

- Occupy in whole or in part lands of the United States.
- Are located on navigable waters in the United States.
- Utilize surplus water or water power from government dams.
- Affect the interests of interstate commerce.

A project's connection to the electrical grid, with transmission lines that cross state boundaries, is considered to be engaged in interstate commerce. Consequently, the vast majority of hydroelectric projects in the United States are subject to FERC jurisdiction.

Often, the first step in the licensing process is to obtain a preliminary permit from FERC. A permit simply reserves the site to the permit holder during the investigation and application phase. Permits are usually issued for a 2-year period, with extension to a third year available. Although a permit provides for a priority advantage in obtaining a fully approved license, under the FPC, municipalities are given preference to hydropower sites.

There is a provision in the licensing process called an exemption. This term may be a misnomer for it does not mean exempt from the licensing process. To achieve an exempt status, an application must still be made. Exemptions are granted to projects meeting certain criteria. One such criterion is that of a project on a man-made conduit, such as at a drop in an irrigation canal. An exemption is granted in perpetuity with no need to apply for an exemption at some future time.

If a site is jurisdictional and ineligible for an exemption, it is necessary to proceed with a formal application. There are two types of licenses. A minor license is for projects under 5 MW, and a major license is for those over 5 MW. Obtaining a license requires a number of different types of studies, consultations with a number of different agencies, preparation of a license application, and can be expensive and time consuming. An FERC license conveys to the license holder the right of eminent domain. Licenses are issued for a specified period of time, usually ranging from 30 to 50 years. Typically, new projects are issued 50-year licenses to offset major capital investments into the project. Any significant change to a project, particularly one affecting the aquatic environment, requires a reopening of the existing license. A change in generating capacity that uses more or less water can have an effect on the aquatic environment.

9.7 UNIQUE FEATURES AND BENEFITS OF HYDRO

9.7.1 Water Resources

Hydroelectric power generation is only one of several potential benefits of river resources development. Multipurpose hydropower projects also provide flood control, flow augmentation, irrigation, municipal water supply, navigation, and recreation opportunities. Hydropower plants convert about 90% of the energy in falling water into electric energy. This is much more efficient than fossil-fueled powerplants, which lose more than half of the energy content of their fuel as waste heat and gases. Hydropower is free of fossil fuel emissions and does not contribute to air pollution, acid rain, or global warming. Furthermore, no trucks, trains, barges, or pipelines are needed to bring fuel to the powerplant site. The earth's hydrologic cycle provides a continual supply of water from rainfall and snowmelt, making hydropower one of the most economic energy resources. And because hydropower is especially suited for providing peaking power, hydroelectricity complements thermal generation and improves overall power production efficiency. Hydro resources often allow utilities to delay or forego construction of additional peaking capacity.

9.7.2 Ancillary Services

Ancillary services comprise the resources and functions (excluding basic generation and transmission capacity) required to support the transfer of electrical energy from generating sources to loads while maintaining reliable operation of the interconnected transmission system. There are several critical ancillary services, which hydro generators are especially effective in providing. These services include the following:

Reactive Supply and Voltage Control. The provision of reactive power from generation sources to support transmission system operations, including the ability to continually adjust transmission system voltage in response to system changes. This service is required to maintain voltage control and stability. Hydroelectric generators, operating in synchronous condense mode, are capable of producing reactive power up to the nameplate capacity of the unit.

Regulation. The provision of adequate generation response capability. Under automatic generation control, supply resources are continuously balanced with minute-to-minute load variations. This service is required to maintain frequency at scheduled values and to help ensure that instantaneous tie line deviations do not cause degradation of transmission system reliability.

Spinning Reserve. Generation capacity is synchronized to the system but is unloaded and able to respond immediately to serve load in case of a system contingency. Capacity is fully available within 10 minutes.

Black Start. The ability of a generating unit or station, during a system restoration, to go from a complete shutdown condition to an operating condition and start delivering power without assistance from the electric system. Requires a dedicated power source for auxiliary equipment and the ability to create own field in exciter. Only required in areas that may become isolated.

9.7.3 Pumped Storage

Pumped-storage plants differ from conventional hydroelectric projects. In a pumped storage scheme, the power station is located between an upper and a lower dam. During periods of high electrical demand, the plant is operated in generating mode. Water is released from the upper dam through the station's turbines and into the lower dam where it is stored. During periods when demand for electricity is low, the machines are put into pump mode to pump water from the lower dam back into the upper dam where it is stored until the station needs to generate again. Pumped storage schemes are net consumers of electricity.

Early pumped storage projects involved separate pumps and turbines. Since the economics of pumped storage favor the highest possible head, configurations included both single and multiple stage pumps and turbines. Sometimes separate motors and generators and even separate penstocks were used. Eventually, reversible pump turbines, in which the pump and turbine are the same machine, were developed. These are not turbines, but are actually pumps with centrifugal impellors, that when operated in the reverse rotational direction are capable of generating as turbines.

The design process of selecting a pump turbine is similar to that of a conventional turbine. One factor of note is that specific speed for a pump is calculated from a different formula, $N_s = N(Q)^{1/2}/H^{3/4}$, where in U.S. units N_s is specific speed, N is rotational speed in rpm, Q is pump discharge in gallons per minute, and H is total dynamic head in feet.

It is an inherent characteristic of reversible pump turbines that the peak efficiency in the generating mode occurs at a slower rotational speed than in the pumping mode. Therefore, unless a more costly 2-speed motor-generator can be used, the selected single, synchronous speed is a compromise speed. Thus, neither the turbine nor pump can operate at their individual peak efficiency. Another factor in this compromise selection of synchronous speed is that the turbine peak efficiency is usually higher than the pump peak efficiency. This is because a pump has additional internal losses including "recirculation" losses. Additionally, the shape of the efficiency profile for the pump mode is more sharply peaked than for the turbine mode.

Generally, a pump turbine needs to spend more time out of a 24-h period in the pumping mode than in the generating mode for the same water exchange. This ratio is referred to as the *duty cycle.* For example, if a pumped storage project pumps for 16 h in order to generate at rated capacity for the remaining 8 h, it is said to have a 16-h duty cycle.

9.8 ENVIRONMENTAL CONCERNS

9.8.1 Fish Passage

Addressing the environmental impact on rivers and maintaining a balance with the plants, fish and wildlife that also depend on the river has never been more difficult. Depending on the particular site of a hydroelectric project, providing for the passage of upstream and downstream migrants can be an important factor. While it may not be a factor for such sites as drops in man-made irrigation canals, it is a critical factor at sites on rivers with anadromous and catadromous fish. Anadromous denotes fish, such as salmon, that mature at sea, but return to fresh water to reproduce. Catadromous are the opposite in that they mature in fresh water, but reproduce at sea, such as certain types of eels. Hydroelectric projects at such migration sites are usually required to be designed with specific upstream and downstream passage facilities.

There are several different types of passage facilities for upstream migrants. The most common is a fish ladder. This is an open surface, shallow gradient, conduit with a series of small plunge pools to dissipate the hydraulic energy from the forebay to the tailwater. However, there are different designs of fish ladders for different migrant species. Salmon will climb the ladder by jumping the weirs connecting each plunge pool. However, shad will not jump, but will climb ladders that have small holes cut in the each weir at the bottom of the each plunge pool. Often extra water is released directly from the forebay as a submerged jet at the entrance to the fish ladder. This is referred to as "fish attraction water." At some powerhouses, overflow weirs are constructed along the entire downstream length for fish to migrate into and be channeled into the fish ladders. Another design of an upstream migrant facility is a fish elevator. In this design, a "crowder" is used to gather the fish into an elevator bucket that is hoisted up to the forebay elevation.

For downstream migrants, there are also a number of design options. One of the oldest is to simply open the gates on controlled spillways during the migratory season as a "fish flush." Another is to collect the downstream migrants, such as at fish hatcheries, and transport them by barges around the hydroelectric projects. Still another option is to install fish screens in the intakes to divert fish from going through the turbines, but into channels that carry them around the powerhouse. There are several different types of fish screens, including submerged bar, extended submerged bar, traveling, etc. Fish screens do disrupt the hydraulics within a turbine and decrease its overall efficiency. A unique type of fish screen called the Eicher fish screen is a wedge wire screen installed downstream of the intake in a penstock. Named after the inventor, George Eicher, it is a self-cleaning screen. This self-cleaning is achieved by simply tilting the screen on horizontal pinions so that the downstream side faces upstream.

Downstream migrants most commonly encounter lower head projects with propeller turbines. For the downstream migrants that do pass through the turbine water passages, five mechanisms of injury have been categorized and studied: strike, grinding and abrasion, decompression, shear and turbulence, and cavitation. For properly submerged lower head projects with propeller turbines, the latter three causes of injury are not as important as the first two. Consequently, to minimize strike, and grinding and abrasion, a unique design of Kaplan runner called a minimum gap runner has been developed. In this design of runner, overhangs and recesses in the runner hub hydraulically hide gaps between the inner edge of the blades and hub, and spherical rather than cylindrical cavities form the discharge ring.

9.8.2 Water Temperature

The presence of a hydroelectric project can improve or degrade the temperature of the aquatic environment. The increased cross section of a forebay or upstream reservoir of a project acts to slow the velocity of a natural river. This tends to decrease the mixing of the vertical water column and leads to stratification. Limnologists, who scientifically study bodies of fresh water, classify this stratification into three distinct layers. Upper most is the epilimnion, where the water is warmed by sunlight. Then at a depth, where sunlight no longer penetrates, there is a thermocline, where the temperature drops rapidly. Below that is the layer of cooler water called the hypolimnion. During the summer months these layers tend to be well defined. However, in winter this distinction tends to fade and may even reverse with the upper layer becoming the coolest and sinking to the bottom.

Hydroelectric projects can be designed to mitigate this temperature stratification or even improve the downstream effect of natural impoundments by two basic approaches. The first is to cause the mixing of the vertical water column. Bubbler hoses can be placed along the bottom of the reservoir to develop air curtains that set up vertical circulation patterns. Large, unhoused, mixing propellers can be used in a similar manner. Fountain-like aerators can be used to spray water from near the bottom of the reservoir into the air.

The second basic approach is to selectively withdraw water from different elevations in the reservoir into the powerhouse intakes. One way this is done is by designing thermal withdrawal towers with foundations that rest on the bottom of the reservoir. These towers have gated ports at different elevations. The water from different reservoir elevations tends to mix inside the tower before entering the penstock. A similar design is to retrofit thermal withdrawal enclosures around conventional powerhouse intakes. These also have gated ports at different elevations and the water from the different elevations tends to mix in the intake. Such structures allow hydroelectric projects to even improve the natural environment, particularly, by discharging cooler water from the bottom of the reservoir in summer.

Aside from the biological factors, water temperature has a minor effect on generation. Temperature, along with elevation and latitude, determines the specific weight of water. The heavier the water, the more electricity that can be generated by a given quantity. Typically, fresh water has a specific weight in U.S. units of about 62.4 lb/ft^3. If this value is divided by the local acceleration of gravity g, in feet per second squared, the value of the density of water is obtained in slugs per cubic foot.

9.8.3 Dissolved Oxygen

Hydroelectric projects can affect the dissolved oxygen (DO) content of their aquatic environment in both beneficial and detrimental ways. Among other things, DO is one of the best indicators of the health of a water ecosystem. In a natural body of water, decrease in the dissolved oxygen levels is often an indication of an influx of some type of organic pollutant. The rate at which oxygen is depleted, usually measured over a 5-day period, is the biological oxygen demand (BOD). Oxygen is consumed by plants and animals during respiration and by aerobic bacteria during the process of decomposition. As a consequence, oxygen consumption is greatest near the bottom of a reservoir, in the hypolimnion, where sunken organic matter accumulates and decomposes. Conversely, oxygen is produced by direct absorption from the atmosphere, by plant photosynthesis or is obtained from inflowing streams. Since photosynthesis requires sunlight, and the air/water interface is at the surface, the higher concentrations of DO are found in the higher elevations of a reservoir, in the epilimnion. Physically, DO can range from 0 to 18 parts per million (ppm), but 5 to 6 ppm are needed to maintain a diverse biota population. DO concentration is most affected by water temperature. Cold water can hold more of any gas than warmer water. The DO concentrations may vary significantly at any time and place due to a number of factors besides temperature, such as barometric pressure, elevation, salinity, season, time of day, wind, and reservoir depth.

Another atmospheric gas that can be of concern is nitrogen. Unlike oxygen, nitrogen is biologically inert. However, it is capable of existing in a supersaturated state for a long period of time. If the gas is in sufficient excess, it can cause a potentially lethal condition, known as “fish bubble disease.” This condition, similar to the bends in divers, is caused when nitrogen comes out of solution in the tissues of fish. Rivers can become supersaturated with nitrogen where the hydraulic jump at the base of spillways entraps large amounts of air in deep pools. If these deep strata are not exposed again to the atmosphere, such that the nitrogen can be "cleansed," it will go into solution. This tends to happen when the tailwater is not free flowing, but is the reservoir of the next downstream hydroelectric project.

Although a powerhouse can do little relative to preventing oxygen depletion in the reservoir or nitrogen supersaturation in the tailrace, it can improve the environment by significantly increasing the DO level of the discharge. The most common method is by aspiration of atmospheric air immediately downstream of the runner of a reaction turbine. The area under the runner at the start of the draft tube is at a low pressure, usually subatmospheric. This is by design to maximize the pressure differential across the runner in order to maximize efficiency. However, by installing air pipes or using existing piping, such as vacuum breakers, atmospheric air can be drawn in or aspirated into that area. Some of the atmospheric oxygen will go into solution and increase the DO of the discharge. The amount of DO increase depends on the water temperature, the ratio of flow rates of air to water, the size of the air bubbles, the distribution of the air in the cross section of hydraulic flow, and the contact time before the bubbles are vented to the surface of the tailwater. This aspiration process does increase the pressure under the runner and this decreases generating efficiency.

Another method to increase DO is to inject commercial oxygen into the powerhouse intake. This is done only rarely because of the expense. Still another method is to use air bubbler hoses laid along the bottom of the reservoir.

BIBLIOGRAPHY

ASME, *Joint ASME-CSME Applied Mechanics, Fluids Engineering, and Bioengineering Conference,* Niagara Falls, New York, American Society of Civil Engineers, 1979.

ASME, *Compendium of Pumped Storage Plants in the United States,* New York American Society of Civil Engineers, 1993.

ASME, *Hydroelectric Pumped Storage Technology,* New York, American Society of Civil Engineers, 1996.

Blank, Z., *Future for Energy Storage Systems,* Stamford, Conn., Business Communications Company, 1975.

Bureau of Reclamation, "Glossary of Hydropower Terms," http://www.usbr.gov/power/edu/edu.html.

Butler, J. G., *How to Build and Operate Your Own Small Hydroelectric Plant,* New York, McGraw-Hill, 1982.

Esposito, A., *Fluid Mechanics with Applications,* New York, McGraw-Hill, 1997.

Foundation for Water and Energy Education, "About Hydropower," http://www.fwee.org/abhydro.html.

Foundation for Water and Energy Education, "Glossary of Terms, Agencies, Laws, and Organizations," http://www.fwee.org/h-glossary.html.

Frankena, F., *Directories for Small-Scale Hydropower Development,* Monticello, Ill., Vance Bibliographies, 1985.

Goodman, L. J., Hawkins, J. N., and Love, R. N., eds., *Small Hydroelectric Projects for Rural Development,* New York, Pergamon Press, 1981.

Granger, R. A., *Fluid Mechanics,* New York, McGraw-Hill, 1995.

Gulliver, J. S., and Arndt, R. E. A., *Hydropower Engineering Handbook,* New York, McGraw-Hill, 1991.

Inversin, A. R., *Micro-hydropower Sourcebook: A Practical Guide to Design and Implementation in Developing Countries,* Washington, D.C., NRECA International Foundation, 1986.

Jog, M. G., *Hydro-electric and Pumped Storage Plants,* New York, Wiley, 1989.

Lockerby, R. W., *Low-Head, Small-Hydro Power Generation,* Monticello, Ill., Vance Bibliographies, 1981.

Morris, G. L., *Reservoir Sedimentation Handbook: Design and Management of Dams, Reservoirs, and Watersheds for Sustainable Use,* New York, McGraw-Hill, 1997.

Munson, B., *Fundamentals of Fluid Mechanics,* New York, McGraw-Hill, 1997.

Roberson, J. A., and Crowe, C. T., *Engineering Fluid Mechanics,* New York, McGraw-Hill, 1996.

Roberts, D. J., *Water-Resources Reports,* Denver, Colo., USGS Earth Science Information Center, Open-File Reports Section, U.S. Geological Survey, 1995.

Simmons, G. M., and Neff, S. E., *The Effect of Pumped-Storage Reservoir Operation on Biological Productivity and Water Quality,* Blacksburg, Va., Water Resources Research Center, Virginia Polytechnic Institute, 1969.

Stage, S., *Control of the Economic Loading of a Large Hydro and Thermal Power Plant,* Stockholm, Vattendraftforeningen, 1955.

Streeter, V. L., *Fluid Mechanics,* New York, McGraw-Hill, 1998.

Tokaty, G. A., *A History and Philosophy of Fluid Mechanics,* New York, McGraw-Hill, 1994.

United Nations, *Mini-Hydropower Stations: A Manual for Decision Makers,* prepared in cooperation with the Latin American Energy Organization (OLADE), New York, United Nations, 1983.

U.S. Army Corps of Engineers Engineering Manual, "Planning and Design of Hydroelectric Power Plant Structures," EM 110-2-3001, 1995.

Vischer, D. L., and Hager, W. H., *Dam Hydraulics,* New York, Wiley Series in Water Resources Engineering, 1987.

Warnick, C. C., Mayo, H. A., Carson, J. L., and Sheldon, L. H., *Hydropower Engineering*, New Jersey, Prentice-Hall, 1984.

Waterpower '87, Conference sponsored by the U.S. Corps of Engineers and Bonneville Power Administration in Portland, Ore., New York, American Society of Civil Engineers, 1988.

Waterpower '95, Proceedings of the International Conference on Hydropower, John J. Cassidy, ed., San Francisco, American Society of Civil Engineers, 1995.

WRP 1992, International Symposium on Hydrology and Water Resources Education and Training: The Challenges to Meet at the Turn of the XXI Century, (Chihuahua, Chihuahua, Mexico), Littleton, Colo., Water Resources Publications, 1992.

SECTION 10
POWER SYSTEM COMPONENTS

Craig A. Colopy
Global Product Manager, Voltage Regulators, Cooper Power Systems

Jon Hilgenkamp
Marketing Manager, Switchgear Products Division, S&C Electric Company

David S. Johnson
President, Pennsylvania Breaker LLC

Robert L. Kleeb
Vice President, ABB Power T&D Company, Inc.

Jeffrey H. Nelson
Principal Electrical Engineer, Substation Projects, Tennessee Valley Authority

Ted W. Olsen
Manager, Technology, Distribution Products Division, Siemens Power T&D

Michael W. Wactor
Senior Design Engineer, R&D Department, Powell Electrical Manufacturing Company

CONTENTS

10.1 TRANSFORMERS

10.1.1 Transformer Theory

Elementary theory is developed from the viewpoint of a 3-phase three-leg concentric-cylindrical two-winding transformer, with the primary low-voltage winding next to the core and the secondary high-voltage winding outside the primary winding. This corresponds to a generator-step-up transformer of moderate kVA. Most of the information is also applicable to single-phase transformers with windings on two legs, 3-phase transformers with five-leg cores, transformers with the primary winding outside the secondary winding, three-winding transformers, substation transformers, etc.

Sinusoidal voltage is induced in windings by sinusoidal variation of flux

$$E = 4.44 \times 10^{-8} a_c BfN \tag{10-1}$$

where a_c = square inches cross section of core, B = lines per square inch peak flux density, E = rms volts, f = frequency in hertz, and N = number of turns in winding.

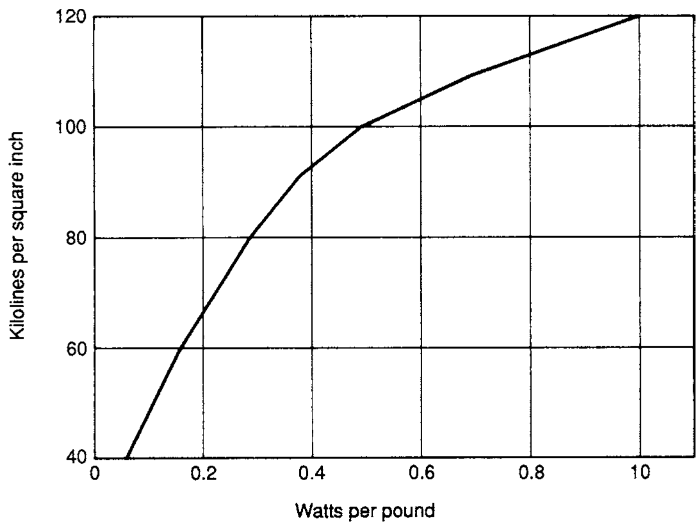


FIGURE 10-1 Typical core-loss curve for transformer core steel at 60 Hz.

The induced voltage in the primary (excited) winding approximately balances the applied voltage. The induced voltage in the secondary (loaded) winding approximately supplies the terminal voltage for the load.

Voltage ratio is the ratio of number of turns ("turn ratio") in the respective windings. The rated open-circuit (no-load) terminal voltages are proportional to the turns in the windings, but under load the primary voltage usually must be somewhat higher than the rated value if rated secondary voltage is to be maintained, because of regulation effects.

Characteristics on Open Circuit. The core loss (no-load loss) of a power transformer may be obtained from an empirical design curve of watts per pound of core steel (Fig. 10-1). Such curves are established by plotting data obtained from transformers of similar construction. The basic loss level is determined by the grade of core steel used and is further influenced by the number and type of joints employed in construction of the core. Figure 10-1 applies for 9-mil-thick M-3-grade steel in a single-phase core with 45° mitered joints. Loss for the same grade of steel in a 3-phase core would usually be 5% to 10% higher.

Exciting current for a power transformer may be established from a similar empirical curve of exciting volt-amperes per pound of core steel as given in Fig. 10-2. The steel grade and core construction are the same as for Fig. 10-1. The exciting current characteristic is influenced primarily by the number, type, and quality of the core joints, and only secondarily by the grade of steel. Because of the more complex joints in the 3-phase core, the exciting volt-amperes will be approximately 50% higher than for the single-phase core.

The exciting current of a transformer contains many harmonic components because of the greatly varying permeability of the steel. For most purposes, it is satisfactory to neglect the harmonics and assume a sinusoidal exciting current of the same effective value. This current may be regarded as composed of a core-loss component in phase with the induced voltage (90° ahead of the flux) and a magnetizing component in phase with the flux, as shown in Fig. 10-3.

Sometimes it is necessary to consider the harmonics of exciting current to avoid inductive interference with communication circuits. The harmonic content of the exciting current increases as the peak flux density is increased. Performance can be predicted by comparison with test data from previous designs using similar core steel and similar construction.

The largest harmonic component of the exciting current is the third. Higher-order harmonics are progressively smaller. For balanced 3-phase transformer banks, the third harmonic components

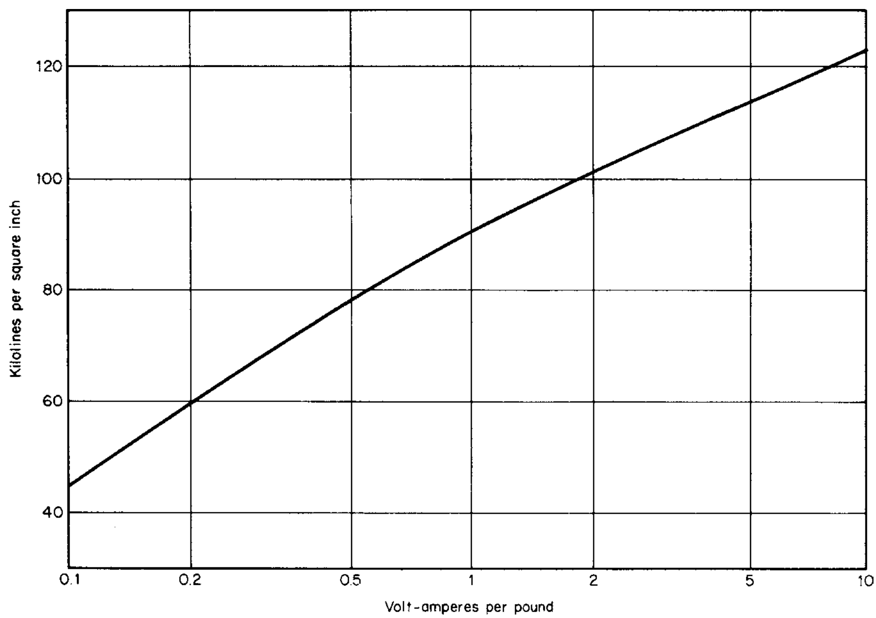


FIGURE 10-2 Typical exciting voltampere curve for transformer core steel at 60 Hz.

(or multiples of the third) are displaced by 120 fundamental degrees (deg) (or multiples of 120 fundamental deg) or 360 harmonic deg and therefore constitute a zero-phase-sequence system. Triple-harmonic currents may flow internally in delta-connected windings and externally in zero phase sequence paths in the connected system. The division of third-harmonic exciting current among available paths is not readily calculable.

Magnetizing Inrush Current. If an idle transformer is energized at a time in the voltage cycle when the flux in the core would normally be other than the actual residual flux in the core, the sinusoidal flux curve will be initially offset, and the offset decreases gradually with time [see Specht (1969) in References list at end of Sec. 10.1.3]. In extreme cases, the peak flux may be more than doubled, exceeding saturation of the core, and causing peak magnetizing current several times rated load current. Magnetizing inrush current is important, principally because of the possibility of false operation of transformer protective relays.

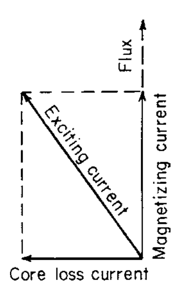


FIGURE 10-3 Phasor diagram of equivalent sinusoidal exciting current.

Characteristics on Short Circuit. If the primary winding of a transformer with 1:1 turn ratio is excited with the secondary winding short-circuited, a small exciting current flows in the primary winding, producing mutual flux mostly in the core. In addition, a short-circuit current flows forward in the primary and reverses in the secondary, causing leakage flux that passes between the two windings and completes its path through the core. The mutual and leakage flux together make net flux linkages with the secondary to induce voltage to supply the resistance drop in the secondary and make net flux linkages with the primary to induce a counter voltage equal to the applied voltage less the resistance drop in the primary. Figure 10-4 shows the space and phase relationships neglecting the exciting current. It is apparent that

$$E_P = I_P(R_P + R_S + jX) = I_P Z \tag{10-2}$$

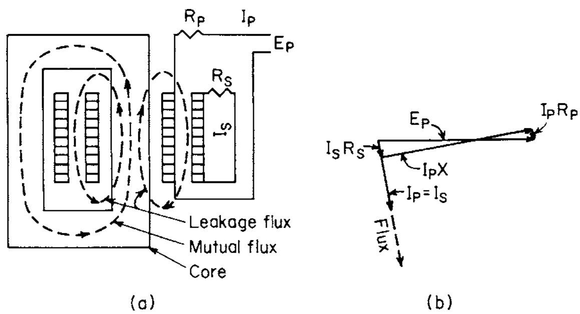


FIGURE 10-4 Short-circuited transformer: (*a*) flux distribution, single-phase; (*b*) phasor diagram, 1:1 ratio.

where E_P = rms volts applied to primary (phasor), I_P = rms amperes in primary (phasor), R_P = ohms resistance of primary winding, R_S = ohms resistance of secondary winding, X = ohms reactance (corresponding to the voltage induced in the primary by the leakage flux), and Z = ohms impedance $(R_P + R_S + jX)$.

Resistance, Reactance, and Impedance. R_P and R_S are effective ac resistances. They are greater than the dc resistances as measured with direct current, because they include eddy loss in the conductor and stray loss in the core clamps, tank, etc. The reactance of the transformer is X, and the impedance is $Z = R_P + R_S + jX$.

Load Loss. The loss on short-circuit test at rated current is the load loss at rated kVA

$$L_L = I_R^2 R = I_R^2 Z_M \cos\theta \tag{10-3}$$

where I_R = rms amperes rated current, L_L = watts load loss at rated current, R = ohms ac resistance $(R_P + R_S)$, Z_M = ohms impedance magnitude $[(R^2 + X^2)^{1/2}]$, and θ = impedance angle of transformer.

The load loss at another current is

$$L = \frac{L_L I^2}{I_R^2} \tag{10-4}$$

where I = rms amperes and L = watts load loss.

Characteristics under Load. Exciting current in the primary winding produces mutual flux mostly in the core. Opposing currents in the primary and secondary windings cause leakage flux, which passes between the two windings and completes its path through the core. The magnitude and phase of the mutual flux depend on the voltage. The magnitude and phase of the leakage flux depend on the current. The mutual and leakage flux together generate in the primary a counter voltage equal to the applied voltage less the resistance drop in the primary, and generate in the secondary a voltage equal to the terminal voltage plus the resistance drop in the secondary.

For most purposes the effect of the leakage flux can be represented by the effect of series reactance in the secondary-winding circuit. Figure 10-5 shows the space relationships and the phase relationships in a transformer of 1:1 ratio. It is apparent that

$$E_P = E_S + I_S(R_S + jX) + I_P R_P \tag{10-5}$$

where E_P = rms volts at primary terminal (phasor), E_S = rms volts at secondary terminal (phasor), I_P = rms amperes in secondary (phasor), I_S = rms amperes in secondary (phasor), R_P = ohms

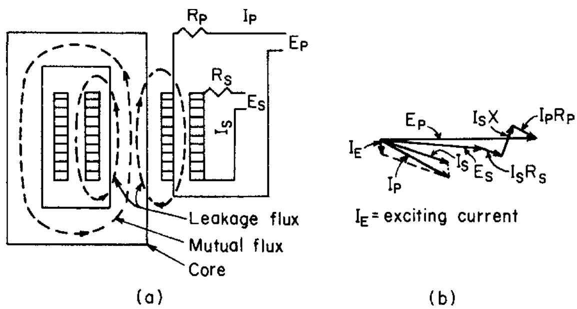


FIGURE 10-5 Loaded transformer: (*a*) flux distribution, single-phase; (*b*) phasor diagram, 1:1 ratio.

resistance of primary winding, R_S = ohms resistance of secondary winding, and X = ohms reactance of transformer.

Equivalent Circuits. Figure 10-6 shows a circuit which for most practical purposes is equivalent to the transformer of Fig. 10-5. The exciting current, I_E, is made up of two components, a magnetizing component flowing through X_M (the major component), and a loss component flowing through R_M. The values of R_M and X_M can be related to Figs. 10-1 and 10-2 if the core flux density at rated voltage is known. It will be found that these quantities vary with the voltage applied to the primary winding and they are usually determined for the rated voltage condition. For many purposes, the exciting current can be neglected and this leads to the simpler circuit of Fig. 10-7.

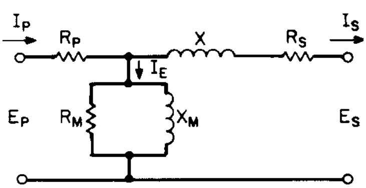


FIGURE 10-6 Equivalent circuit of a two-winding transformer considering exciting current.

Effect of Turn Ratio. Equation (10-5) and Fig. 10-7 represent a transformer of 1:1 turn ratio. A transformer of turn ratio T secondary to primary can be transformed into an

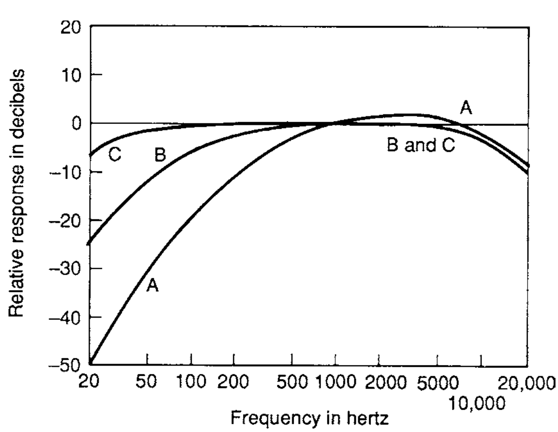


FIGURE 10-7 Equivalent circuit of a two-winding transformer neglecting exciting current.

equivalent 1:1 transformer by imagining the secondary winding replaced by a winding with the same number of turns as the primary winding, but using the same weight of conductor and occupying the same space as the secondary winding. I_S, E_S, and R_S in the real secondary winding become I_S/T, E_S/T, and R_S/T^2. The impedance of the load, Z_L, becomes Z_L/T^2. Thus, although Eqs. (10-2) to (10-5) and Figs. 10-4 to 10-7 were given for 1:1 turn ratio, they can be applied to any turn ratio. The fact that the simple series impedance of Fig. 10-7 may be used as equivalent to a transformer of any turn ratio is very helpful in the analysis of electric power systems. Secondary-winding characteristics corresponding to a fictitious secondary winding of 1:1 turn ratio are called secondary characteristics referred to the primary side. If more convenient, all characteristics can be referred to the secondary side by a reverse process.

Percent and Per Unit. Current, voltage, and kVA are frequently expressed as per unit or percent of rated value (25% = 0.25 per unit). The procedure is extended to resistance, reactance, and impedance by defining per unit impedance as (ohms impedance) × (rated current in amperes) ÷ (rated voltage in volts). Quantities expressed in percent or per unit are the same regardless of whether they are referred to the primary side or the secondary side.

Regulation. It is apparent from Eq. (10-5) that if the load current and the secondary voltage are at rated value, the primary voltage must exceed rated value. The excess is called regulation. Regulation in per unit is defined as the difference between primary and secondary voltage divided by secondary voltage. For rated load at lagging power factor and rated secondary voltage, regulation is given exactly by Eq. (10-6) or approximately by Eq. (10-7).

$$G_r = [(R_r + P_r)^2 + (X_r + Q_r)^2]^{1/2} - 1 \tag{10-6}$$

$$G_0 = 100\left[P_rR_r + Q_rX_r + \frac{(P_rX_r - Q_rR_r)^2}{2}\right] \tag{10-7}$$

where G_0 = percent regulation, G_r = per unit regulation, P_r = per unit load power factor, $Q_r = (1 - P_r^2)^{1/2}$, R_r = per unit resistance of transformer, and X_r = per unit reactance of transformer.

The calculation of regulation of a three-winding transformer is considerably more complex, depending on the load sharing between the two secondary windings. It will not be treated here.

Impedance Data. Resistance and reactance of transformers tend to follow normal patterns according to the ratings. Figure 10-8 shows resistance in percent (as determined by measurement of load loss on impedance test). Specific units may vary as much as ± 30% depending largely on the evaluation of losses as compared with capital cost. Figure 10-9 shows ranges of reactance in percent. Special designs (transformers with all windings high-voltage, autotransformers, designs with

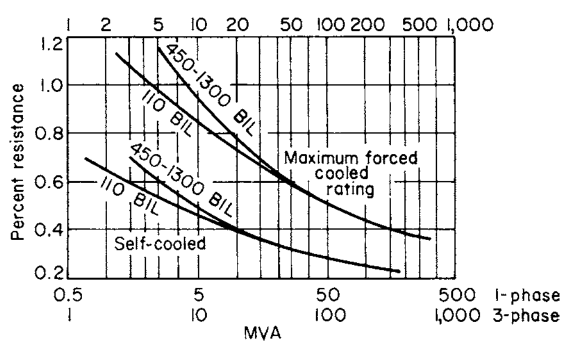


FIGURE 10-8 Resistance of typical power transformer.

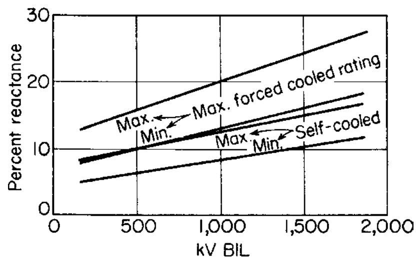


FIGURE 10-9 Reactance of typical power transformer.

overload ratings, etc.) may have reactances outside the limits shown.

Efficiency. This is given by

$$F_r = \frac{E_S I_S \cos\theta}{E_S I_S \cos\theta + L_{NS} + L_{LS}} \tag{10-8}$$

where E_S = rms volts at secondary terminals, F_r = per unit efficiency, I_S = rms amperes in secondary, L_{LS} = watts load loss at I_S, L_{NS} = watts no-load loss at $E_S(I_S = 0)$, and θ = impedance angle of load.

Three-Winding-Transformer Load Losses. The load losses of three-winding transformers, with all three windings carrying loads simultaneously, may be calculated from characteristics obtained by considering each pair of windings as a two-winding transformer.

$$L_T = \left(\frac{I_P}{I_A}\right)^2 \frac{L_{PS} + L_{PT} - L_{ST}}{2} + \left(\frac{I_{SP}}{I_A}\right)^2 \frac{L_{ST} + L_{PS} - L_{PT}}{2} + \left(\frac{I_{TP}}{I_A}\right)^2 \frac{L_{PT} + L_{ST} - L_{PS}}{2} \tag{10-9}$$

where I_A = rms amperes reference current referred to winding P; I_P = rms amperes in winding P; I_{SP} = rms amperes in winding S referred to winding P; I_{TP} = rms amperes in winding T referred to winding P; L_{PS} = watts load loss in windings P and S as a two-winding transformer at I_A; L_{PT}, L_{ST} = similar; and L_T = watts total load loss.

The loss is usually computed at, or corrected to, a temperature of 75°C for 55°C average rise units and 85°C for 65°C average rise units.

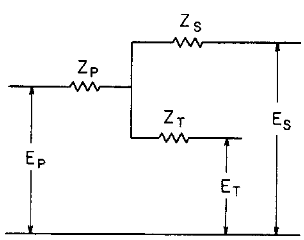


FIGURE 10-10 Equivalent circuit of a three-winding transformer.

Three-Winding-Transformer Equivalent Circuit. The equivalent circuit of a three-winding transformer may be determined from the three impedances obtained by considering each pair of windings separately. One form is shown in Fig. 10-10, in which

$$Z_P = \frac{Z_{PS} + Z_{PT} - Z_{ST}}{2} \tag{10-10}$$

$$Z_S = \frac{Z_{PS} + Z_{ST} - Z_{PT}}{2} \tag{10-11}$$

$$Z_T = \frac{Z_{PT} + Z_{ST} - Z_{PS}}{2} \tag{10-12}$$

where Z_P, Z_S, Z_T = ohms branch impedances in Fig. 10-10; Z_{PS} = ohms impedance from winding P to winding S in two-winding equivalent circuit of Fig. 10-7; and Z_{PT}, Z_{ST} = similar. All ohmic values of impedance must be referred to one common winding (i.e., the primary winding).

Four-Winding-Transformer Equivalent Circuit. The equivalent circuit of a four-winding transformer may be determined from the six impedances obtained by considering each pair of windings separately. One form is shown in Fig. 10-11, in which

$$Z_P = \frac{Z_{PQ} + Z_{PS} - Z_{SQ}}{2} - \frac{Z_A Z_B}{2(Z_A + Z_B)} \tag{10-13}$$

$$Z_S = \frac{Z_{PS} + Z_{ST} - Z_{PT}}{2} - \frac{Z_A Z_B}{2(Z_A + Z_B)} \tag{10-14}$$

$$Z_T = \frac{Z_{ST} + Z_{TQ} - Z_{SQ}}{2} - \frac{Z_A Z_B}{2(Z_A + Z_B)} \tag{10-15}$$

$$Z_Q = \frac{Z_{TQ} + Z_{PQ} - Z_{PT}}{2} - \frac{Z_A Z_B}{2(Z_A + Z_B)} \tag{10-16}$$

$$Z_A = (K_1 K_2)^{1/2} + K_1 \qquad Z_B = (K_1 K_2)^{1/2} + K_2 \tag{10-17}$$

$$K_1 = Z_{PT} + Z_{SQ} - Z_{PS} - Z_{TQ} \qquad K_2 = Z_{PT} + Z_{SQ} - Z_{PQ} - Z_{ST} \tag{10-18}$$

where Z_A = ohms branch impedance in Fig. 10-11 (complex); Z_B, Z_P, Z_S, Z_T, Z_Q = similar; Z_{PS} = ohms impedance (complex) from winding P to winding S in two-winding equivalent circuit of Fig. 10-7; and Z_{PT}, Z_{PQ}, Z_{ST}, Z_{SQ}, Z_{TQ} = similar.

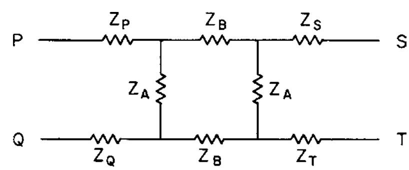


FIGURE 10-11 Equivalent circuit of a four-winding transformer.

Phase-Interconnected Transformers. These transformers that is with windings from more than one phase on a single core leg can be represented by an equivalent circuit only if each winding on a leg is considered as if it were brought out to separate terminals [see Cogbill (1955) in list at end of Sec. 10.1.3].

10.1.2 Transformer Connections

Parallel Operation. Two single-phase transformers will operate in parallel if they are connected with the same polarity. Two 3-phase transformers will operate in parallel if they have the same winding arrangement (e.g. Y-delta), are connected with the same polarity, and have the same phase rotation. If two transformers (or two banks of transformers) have the same voltage ratings, the same turn ratios, the same impedances (in percent), and the same ratios of reactance to resistance, they will divide the load current in proportion to their kVa ratings, with no phase difference between the currents in the two transformers. If any of the above conditions are not met, the load current may not divide between the two transformers in proportion to their kVA ratings and there may be a phase difference between currents in the two transformers.

Two unlike transformers connected in parallel will supply current to a load as follows:

$$I_L = \frac{E_P}{\dfrac{1}{(T_1/Z_1) + (T_2/Z_2)} + Z_L \dfrac{(1/Z_1) + (1/Z_2)}{(T_1/Z_1) + (T_2/Z_2)}} \tag{10-19}$$

where E_P = rms volts on primary side (phasor), I_L = rms amperes total load current (phasor), T_1 = turn ratio secondary to primary of unit 1, T_2 = turn ratio secondary to primary of unit 2, Z_1 = ohms impedance of unit 1 referred to secondary side (complex), Z_2 = ohms impedance of unit 2 referred to secondary side (complex), and Z_L = ohms impedance of load (complex).

The magnitude of the current in unit 1 is

$$I_{r1} = \frac{\{[T_1 R_{r2} I_{rL} + (T_1 - T_2) E_{r1} \cos\theta]^2 + [T_1 X_{R2} I_{rL} + (T_1 - T_2) E_{r1} \sin\theta]^2\}^{1/2}}{[(T_1 R_{r2} + T_2 R_{r1})^2 + (T_1 X_{r2} + T_2 X_{r1})^2]^{1/2}} \tag{10-20}$$

where E_{r1} = rms voltage of secondary terminals in per unit of unit 1, I_{rL} = rms total load current in per unit of unit 1, I_{r1} = rms current in secondary of unit 1 in per unit of unit 1, T_1 = ratio secondary turns to primary turns in unit 1, T_2 = ratio secondary turns to primary turns in unit 2, R_{r1} = equivalent resistance of unit 1 in per unit of unit 1, R_{r2} = equivalent resistance of unit 2 in per unit of unit 1, X_{r1} = equivalent reactance of unit 1 in per unit of unit 1, X_{r2} = equivalent reactance of unit 2 in per unit of unit 1, and θ = impedance angle of load (lagging current positive).

NOTE: Per unit means percent divided by 100, that is, 10% = 0.1 per unit.

The current in the second unit may be determined by using Eq. (10-20) with designation of first and second transformers reversed.

Phase-interconnected transformers (i.e., with windings from more than one phase on a single core leg) offer special complication when unlike units are connected in parallel. [See Cogbill (1955) in list at end of Sec. 10.1.3.]

3-Phase to 3-Phase Transformations. The delta-delta, the delta-Y, and the Y-Y connections are the most generally used; they are illustrated in Fig. 10-12. The Y-delta and delta-delta connections may be used as step-up transformers for moderate voltages. The Y-delta has the advantage of providing a good grounding point on the Y-connected side which does not shift with unbalanced load and has the further advantage of being free from third-harmonic voltages and currents; the delta-delta has the advantage of permitting operation in V in case of damage to one of the units. Delta connections are not the best for transmission at very high voltage; they may, however, be associated at some point with other connections that provide means for properly grounding the high-voltage system; but it is better, on the whole, to avoid mixed systems of connections. The delta-Y step-up and Y-delta step-down connections are without question the best for high-voltage transmission systems. They are economical in cost, and provide a stable neutral whereby the high-voltage system may be directly grounded or grounded through resistance of such value as to damp the system critically and prevent the possibility of oscillation.

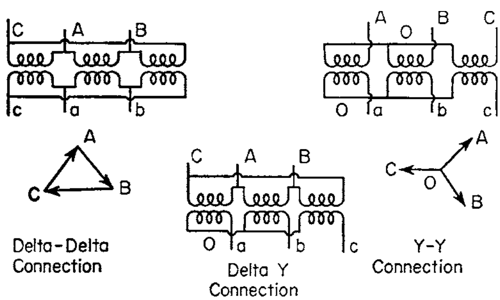


FIGURE 10-12 Standard 3-phase/3-phase transformer systems.

The Y-Y connection (or Y-connected autotransformer) may be used to interconnect two delta systems and provide suitable neutrals for grounding both of them. A Y-connected autotransformer may be used to interconnect two Y systems which already have neutral grounds, for reasons of economy. In either case, a delta-connected tertiary winding is frequently provided for one or more of the following purposes.

In *stabilization of the neutral,* if a Y-connected transformer (or autotransformer) with a delta-connected tertiary is connected to an ungrounded delta system (or poorly grounded Y system), stability of the system neutral is increased. That is, a single-phase short-circuit to ground on the transmission line will cause less drop in voltage on the short-circuited phase and less rise in voltage on the other two phases. A 3-phase three-leg Y-connected transformer without delta tertiary furnishes very little stabilization of the neutral, and the delta tertiary is generally needed. Other Y connections offer no stabilization of the neutral without a delta tertiary. With increased neutral stabilization, the fault current in the neutral on single-phase short circuit is increased, and this may be needed for improved relay protection of the system.

Third-harmonic components of exciting current find a relatively low impedance path in a delta tertiary on a Y-connected transformer, and less of the third-harmonic exciting current appears in the connected transmission lines, where it might cause interference with communication circuits. Failure to provide a path for third-harmonic current in Y-connected 3-phase shell-type transformers or banks of single-phase transformers will result in excessive third-harmonic voltage from line to neutral. The bank of a 3-phase, three-legged core-type Y-connected transformer acts as a delta winding with high impedance to the other windings. As a consequence, there is very little third-harmonic line-to-neutral voltage and a separate delta tertiary is not needed to reduce it.

An external load can be supplied from a delta tertiary. This may include synchronous or static capacitors to improve system operating conditions.

Loading Y-Connected Transformers Line to Neutral. Load can be connected line to neutral only if (1) the source side of the transformer is delta-connected, or (2) the source side is Y-connected with the neutral connected back to the source neutral.

If one of these two conditions is not maintained, the neutral will shift, reducing the voltage of the loaded phase and increasing the voltage of the other phases.

The Open-Delta Connection, or V Connection. This is an unsymmetric connection which is used if one transformer of a bank of three single-phase delta-connected units must be cut out because of failure. It is a connection that is sometimes resorted to as an emergency expedient or used as a temporary measure with the intention of completing the delta when conditions of load warrant the addition of a third unit. If one phase of a 3-phase delta-connected transformer of the shell type should fail, operation may be continued at reduced capacity by short-circuiting the damaged phase; if of the core type, operation may be continued by leaving the damaged phase open-circuited, provided that the windings are still capable of withstanding the voltage stresses. Since full-line currents flow in the windings out of phase with the transformer voltages, the normal capacity of the open-delta bank is reduced to 57.7% of its delta rating.

The T Connection. This uses two transformers, the first called the "main" transformer, connected from line to line; and the second, called the "teaser" transformer, connected from the midpoint of the first to the third line. It requires that the midpoint of both primary and secondary windings be available for connections. It has an advantage over the V connection in being more nearly symmetrical if the proper taps have been provided. As in the case of the V connection, two transformers of a bank of delta-connected transformers, one of which has failed, may be connected in T, and if 10% taps can be used for the teaser transformer, the transformation will be more nearly symmetrical than if the V connection were used. Where T-connected transformers are installed, they may later be changed to delta with the addition of one more transformer and an increase in rating of the bank of 73%. In the T connection (Fig. 10-13) the transformer *AD*, known as the teaser transformer, may be a duplicate of the main transformer so as to be interchangeable with it, and it may or may not be provided with an 86.6% tap. Its rated capacity will then be 15.5% more than actually necessary. The main transformer operates at a power factor of 0.866, and therefore, if the two transformers are duplicates, their total rated capacity will be 15.5% greater than the capacity of the load in kVA, or each transformer must have a rating of 0.577 of the kVA delivered. If the transformers are not interchangeable, the teaser may be reduced to a rating of one-half the kVA delivered.

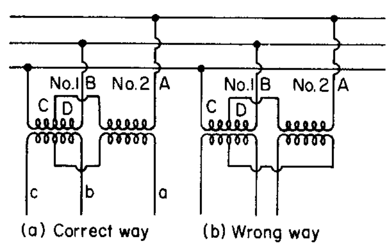


FIGURE 10-13 T-connected transformers.

In connecting transformers in T, *care should be taken to keep the relative phase sequence of the windings the same;* otherwise the impedance of the main transformer may be excessively high and cause undue unbalance. Figure 10-13 illustrates the right and the wrong way.

3-Phase to 6-Phase Transformation. 6 Phases are commonly used for supply of rectifiers. Six phases can be obtained as shown in Fig. 10-14 (double delta) or as in Fig. 10-15 (double Y) respectively. It is not necessary in this transformation, when the neutral connection is required, to have two secondary windings; instead a middle tap may be brought out, all the middle taps being connected together to form the neutral.

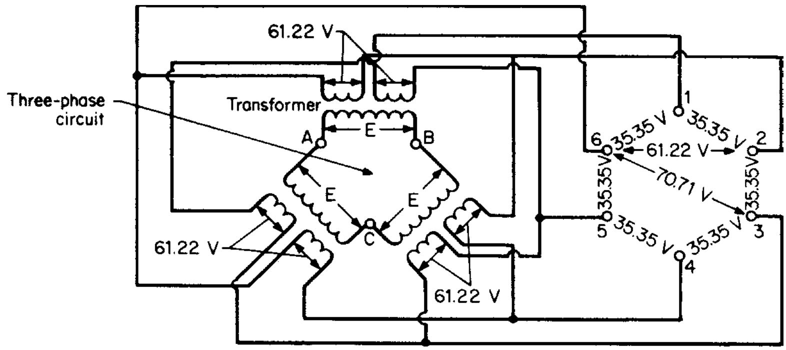


FIGURE 10-14 Three-phase/six-phase transformation, double delta.

The Interconnected Y Connection. This connection (see Fig. 10-16) is commonly referred to as the *zigzag* connection. It may be used with either a delta-connected winding as shown or a Y-connected winding for step-up or step-down operation. In either case, the zigzag winding produces the same angular displacement as a delta winding and, in addition, provides a neutral for grounding purposes. Owing to the angular relation of voltages of the zig and zag windings, the amount of conductor material required for such a connection is 15% greater than a corresponding Y or delta connection. If a transformer consists of zigzag and Y connections, a third winding, delta-connected, is usually necessary for reasons given under the Y-Y connection. If the delta-connected winding is included for purposes other than that of providing a third source of power, in some cases it is practical to design it for the same voltage as the zigzag winding and connect it in parallel with the zigzag winding to form the delta-grounded transformer connection [see Gross and Rao (1953) in list at end of Sec. 10.1.3].

The zigzag connection is used extensively for grounding transformers, the sole purpose of which is to establish a neutral point for grounding purposes; therefore no other windings are required.

10.1.3 Power Transformers

Power transformers may be defined as transformers used to transmit or distribute power in ratings larger than distribution transformers (usually over 500 kVA or over 67 kV). Some of the following information on power transformers is also applicable to other types of transformer.

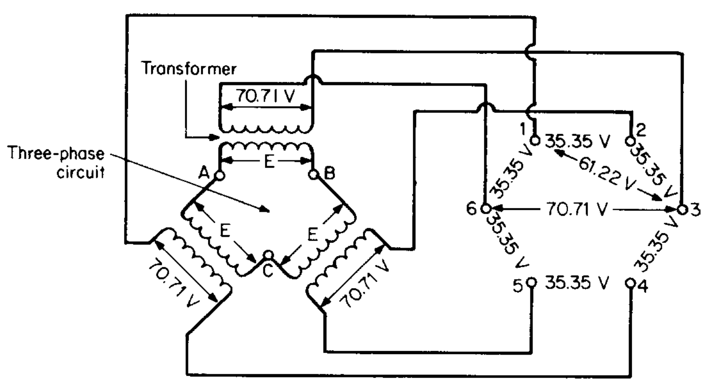


FIGURE 10-15 Three-phase/six-phase transformation, double Y.

The Rated Constants of a Power Transformer. The kVA, terminal voltages and currents are defined in ANSI C57.12.80. They are all based on the rated winding voltages at no load, although it is recognized that the actual primary voltage in service must be higher than the rated value by the amount of the regulation, if the transformer is to deliver rated voltage to the load on the secondary.

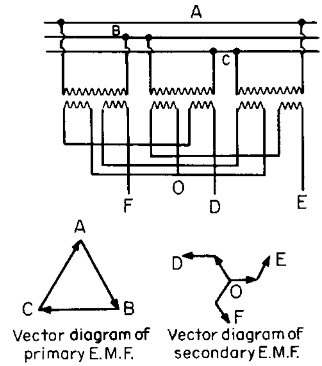


FIGURE 10-16 Interconnected Y connection.

10.1.4 Design

The design of commercial transformers requires the selection of a simple yet suitable form of construction so that the coils are easy to wind and the core is easy to build. At the same time, the mean length of the windings and magnetic circuit must be as short as possible for a given cross-sectional area, so that the amount of material required and resulting losses are minimized. The core must provide a continuous path for magnetic flux while its lamination pattern must be easy to cut and stack. The windings should be insulated in a simple and economical manner, should permit the dissipation of heat (due to losses) by means of cooling ducts, and should be mechanically strong to withstand short-circuit forces.

Two Types of Transformer in Common Use. When the magnetic circuit takes the form of a single ring encircled by two or more groups of primary and secondary windings distributed around the periphery of the ring, the transformer is termed a *core-type transformer*. When the primary and secondary windings take the form of a common ring which is encircled by two or more rings of magnetic material distributed around its periphery, the transformer is termed a *shell-type transformer* (Fig. 10-17). Actually, core-type (or "core-form") in U.S. power-transformer engineering usage means that the coils are cylindrical and concentric (the outer winding over the inner) whereas shell-type (or "form") denotes large pancake coils that are stacked or interleaved to make primary-secondary (P-S) groups. Except for certain extremes of current rating, the choice between the core- and shell-type construction is largely a matter of manufacturing facilities and of individual preference.

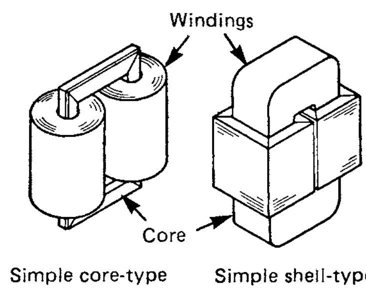


FIGURE 10-17 Forms of magnetic circuits for transformers.

Core-form transformer characteristic features are a long mean length of magnetic circuit and a short mean length of windings. Commonly used core constructions for single-phase and 3-phase units are shown in Figs. 10-18 and 10-19, respectively. The three-leg (one active leg) and four-leg (two active) construction of single-phase cores and the five-leg (three active) construction of 3-phase cores are used to reduce overall height. In these cases, the core encloses the cylindrical windings in

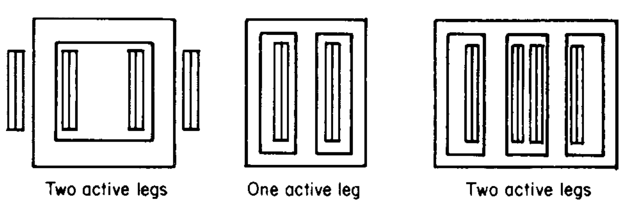


FIGURE 10-18 Single-phase core-form core construction.

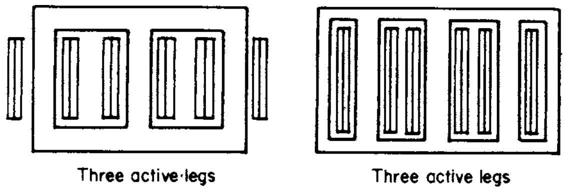


FIGURE 10-19 Three-phase core-form core construction.

a similar fashion to the shell-form construction. The simple concentric primary (inside) and secondary (outside) winding arrangement is common for all small- and medium-power transformers. However, large MVA transformers frequently have some degree of interleaving of windings, such as secondary-primary-secondary (S-P-S). The core-form construction can be used throughout the full size range of power transformers.

Shell-form transformer characteristic features are short mean length of magnetic circuit and long mean length of windings. This results in the shell-form transformer having a larger area of core and smaller number of winding turns than the core form of same output and performance. Also, the shell form would typically have a greater ratio by weight of steel to copper. Figure 10-20 shows the conventional 3-phase shell-form core with the coils in cross section. Primary-secondary-primary (P-S-P) coil grouping is most common, but P-S-P-S-P is often used.

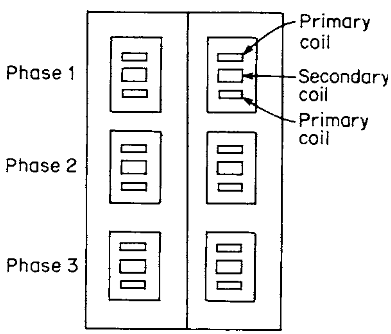


FIGURE 10-20 Conventional 3-phase core for the rectangular-pancake-interleaved-coil structure (shell type). The groups of pancake coils may be round or rectangular.

Design Process. The design process begins with a customer specification. This contains the desired characteristics which allow the design engineer to make economical choices between different materials and construction geometrics. Some of these characteristics are voltage (kV), power rating (MVA), impedance (%), loss evaluation ($/kW), temperature rating (°C), and cooling class (OA, FA, FOA). Some additional factors which influence the transformer design are requirements for tertiary windings and no-load or load taps.

After the customer specification is understood, the design optimization can begin. Most power transformers are designed starting with a certain winding arrangement and dimensions of the core and coils. Initial characteristics of the transformer are calculated and then compared to the desired characteristics. The initial dimensions are then modified to better meet the desired characteristics. Repeating the process leads to close agreement of calculated characteristics with desired characteristics. The repeated calculations, converging on the optimum design, are usually performed by computer. Closeness of agreement of calculated characteristics with tested characteristics depends upon the degree of refinement of the design process, the closeness of agreement of the physical properties of the materials used (particularly the dielectric properties of the insulating materials and the magnetic properties of the core steel) with the properties assumed in the design calculation, and the accuracy of the manufacturing procedures and processes.

Refinement of the design process results from comparison with test data obtained on similar transformers. This applies particularly to core loss, stray loss, noise level, reactance, and dielectric strength. The following calculation methods are mostly approximate.

With an assumed core cross section and flux density, the *number of turns* in each winding is established from Eq. (10-1). The flux density is adjusted to give an integral number of turns in the low-voltage winding, and then an acceptable ratio of open-circuit terminal voltage results from an integral number of turns in the high-voltage winding.

Leakage flux density in the main gap (insulation space between windings) for a transformer with one core leg per phase, as shown in Fig. 10-21, is as follows:

$$B_L = \frac{4.52 I_R N}{h_E} \qquad (10\text{-}21)$$

where B_L = lines per square inch peak leakage flux density, h_E = inches effective length of leakage flux path, I_R = rms amperes rated current of winding, and N = number of turns in winding.

If there is more than one leg per phase, Eq. (10-21) applies to the portion of winding on one leg. The effective length of leakage path is difficult to evaluate accurately. For concentric cylindrical windings, it is approximately

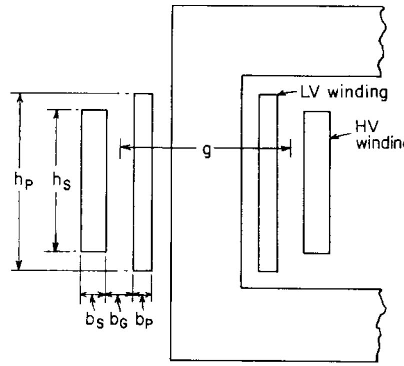


FIGURE 10-21 Dimensions of core-type concentric windings for reactance calculations.

$$h_E = \frac{h_P + h_S}{2} + 0.8\left(\frac{b_P + b_S}{3} + b_G\right) \qquad (10\text{-}22)$$

where b_G = inches radial distance between windings, b_P = inches radial width of winding P, b_S = inches radial width of winding, S, h_P = inches length of winding P, and h_S = inches length of winding S.

Leakage reactance may be calculated for a transformer with one set of coils per phase as follows:

$$X_r = \frac{2.01 \times 10^{-7}\, a_L f I_R N^2}{E_R h_E} \qquad (10\text{-}23)$$

where a_L = square inches effective cross section of leakage flux path, E_R = rms volts rated voltage of winding, f = frequency in hertz, h_F = effective length of leakage flux path in inches, I_R = rms amperes rated current of winding, N = number of turns in winding, and X_r = per unit reactance.

The effective cross section of the leakage flux path is difficult to evaluate accurately. For concentric cylindrical windings, it is approximately

$$a_L = \left(b_G + \frac{b_P + b_S}{3}\right)\pi g \qquad (10\text{-}24)$$

where g = mean diameter of main gap, in.

Resistance loss in winding is

$$W_R = 2.57\, M^2 \frac{234.5 + C}{309.5} \qquad (10\text{-}25)$$

$$L_R = H_C W_R \qquad (10\text{-}26)$$

where L_R = watts resistance loss in winding, M = rms kiloamperes per square inch current density, H_C = pounds weight of copper in winding, C = temperature in degrees Celsius, and W_R = watts per pound resistance loss in winding.

Eddy loss in the winding may be regarded as caused by circulating current induced in the strand by the magnetic flux passing through the strand. For a two-winding transformer,

$$W_E = 2.06 \times 10^{-10} d^2 f^2 B_L^2 \frac{309.5}{234.5 + C} \qquad (10\text{-}27)$$

$$L_E = H_C W_E \qquad (10\text{-}28)$$

where B_L = lines per square inch peak leakage flux density, from Eq. (10-21), C = Celsius temperature, d = inches thickness of strand perpendicular to flux, f = hertz frequency, L_E = watts eddy loss in winding, H_C = pounds weight of copper in winding, and W_E = watts per pound average eddy loss in winding.

Load loss is the sum of resistance and eddy losses in all windings plus stray loss. The stray loss is in itself an eddy loss produced by the leakage flux penetrating the surface of other conducting components, such as the core, core clamps, and tank. Historically, the stray loss has been predicted from test results on similar transformers, but finite element computer solutions have recently been developed to define the leakage flux paths accurately and permit more exact calculation of both stray and eddy losses.

No-load loss equals watts per pound determined from Fig. 10-1, multiplied by weight of the core multiplied by a correction factor depending on core configuration and processing and determined by experience.

General Design Characteristics. The relationship of power-transformer characteristics to scale factor can be illuminated by considering the effect of increasing all dimensions in the ratio S, while retaining the same thickness of core lamination and thickness of conductor strand, but imagining the conductor turns to be reconnected for a terminal voltage proportionate to the insulation thickness. Similarly the effect of increasing the flux density in the ratio B and the current density in the ratio M can be examined. The results are shown in Table 10-1.

Core dimensions are generally standardized in steps, with only a small number of dimensions varying to meet the requirement of the particular rating. Cold-rolled grain-oriented silicon steel strip in gages of 0.009 to 0.014 in is used with mitered corner joints to take advantage of the good characteristics of this material when carrying flux in the rolling direction.

10.1.5 Insulation

Insulation systems in power transformers consist of a fluid—either liquid or gas—together with solid materials. Petroleum-based oils have been used to insulate power transformers since 1886 and are still used in virtually all medium and large transformers (Sheppard 1986). Askeral was used from

TABLE 10-1 Scale Effects

Characteristic	At scale factor S	At flux density B*	At current density M
Linear dimension	S		
Flux density	. . .	B	
Current density	. . .	. . .	M
Rated current	S^3	B	M
Rated voltage	S		
Rated kVA	S^4	B	M
Weight	S^3		
kVA/lb	S	B	M
Ω reactance	S^{-1}		
% reactance	S	B^{-1}	M
W core loss	S^3	B^2	
% core loss	S^{-1}	B	M^{-1}
W $I^2 R$ loss	S^2	. . .	M^2
% $I^2 R$ loss	S^{-1}	B^{-1}	M
W eddy loss	S^5	. . .	M^2
% eddy loss	S	B^{-1}	M
W stray loss†	S^4	. . .	M^2
% stray loss†	. . .	B^{-1}	M

*Applies only to the range in which core loss varies with the square of B.
†As a percent of rated current times rated volts.

1932 through the mid-1970s when the flammability of mineral oil was a concern, but it has since been completely phased out of transformer production because of environmental concerns. It has been replaced by any of a wide variety of high-flash-point fluids (silicones, high-flash-point hydrocarbons, clorinated benzenes, or chlorofluorocarbons). Gas systems include nitrogen, air, and fluorogases. The fluorogases are used to avoid combustability and limit secondary effects of internal failure. Some transformers have been constructed using low boiling-point liquids such as Freon which allows improved heat transfer using a 2-phase cooling system.

Within the core and coil assembly, insulation can be divided into two fundamental groups: major insulation and minor insulation. Major insulation separates the high- and low-voltage windings, and the windings to core. Minor insulation may be used between the parts of individual coils or windings depending on construction. Finally, turn insulation is applied to each strand of conductor and/or groups of strands forming a single turn.

Oil-Insulated Transformers. Low cost, high dielectric strength, excellent heat transfer characteristics, and ability to recover after dielectric overstress make mineral oil the most widely used transformer insulating material. The oil is reinforced with solid insulation in various ways. The major insulation usually includes barriers of wood-based paperboard (pressboard), the barriers usually alternating with oil spaces. Because the dielectric constant of the oil is 2.2 and that of the solid is approximately 4.0, the dielectric stress in the oil ends up being higher than that of the pressboard, and the design of the structure is usually limited by the stress in the oil.

The insulation on the conductors of the winding may be enamel or wrapped paper which is either wood- or nylon-based. The use of insulation directly on the conductor actually inhibits the formation of potentially harmful streamers in the oil, thereby increasing the strength of the structure (Nelson 1989). Again, the limit of dielectric strength is usually that of the oil.

Heavy paper wrapping is also usually used on the leads coming from the winding. In this case, the insulation serves to reduce the stress in the oil by moving the interface from the surface of the conductor (where the stress is high) to a distance away from the conductor (where the stress is considerably lower). Again, the stress in the oil determines the amount of paper required, and the thermal considerations establish the minimum size of the conductor for the necessary insulation.

Askeral-Insulated Transformers. These transformers have constructions similar to the oil-insulated transformers. The relatively high dielectric constant of the askeral aids in transferring the dielectric stress to the solid elements. Askeral has limited ability to recover after dielectric overstress, and thus the strength is limited in nonuniform dielectric fields. Askerals are seldom used over 34.5-kV operating voltage. They are powerful solvents; their products of decomposition are so harmful that they have been completely abandoned in transformers manufactured after the mid-1970s.

Fluorogas-Insulated Transformers. Fluorogases have better dielectric strength than nitrogen or air. Although their heat transfer characteristics are poorer than oil, they are better than nitrogen or air because of their higher density. Both dielectric strength and heat transfer capability increase with pressure; in fact, the dielectric strength at 3 atm gage pressure—where some fluorocarbon-insulated transformers operate—can approach that of oil. The gas insulation is reinforced with solid insulation used in the form of barriers, layer or disk insulation, turn insulation, and lead insulation similar to oil-immersed transformers.

It is usually economical to operate fluorogas-insulated transformers at higher temperatures than oil-insulated transformers. Suitable solid insulating materials include glass, asbestos, mica, high-temperature resins, and ceramics.

Dielectric stress on the gas is several times higher than in the adjacent solid insulation; care must be taken to avoid overstressing the gas.

Nitrogen and Air-Insulated Transformers. These are generally limited to 34.5 kV and lower operating voltages. Air-insulated transformers in clean locations are frequently ventilated to the atmosphere. In contaminated atmospheres a sealed construction is required, and nitrogen is generally used at approximately 1 atm and some elevated operating temperatures.

Design of Insulation Structures. Three factors must be considered in the evaluation of the dielectric capability of an insulation structure—the voltage distribution must be calculated between different parts of the winding, the dielectric stresses are then calculated knowing the voltages and the geometry, and finally the actual stresses can be compared with breakdown or design stresses to determine the design margin.

Voltage distributions are linear when the flux in the core is established. This occurs during all power frequency test and operating conditions and to a great extent under switching impulse conditions. (Switching impulse waves have front times in the order of tens to hundreds of microseconds and tails in excess of 1000 μs.) These conditions tend to stress the major insulation and not inside of the winding. For shorter-duration impulses, such as full-wave, chopped-wave, or front-wave, the voltage does not divide linearly within the winding and must be determined by calculation or low-voltage measurement. The initial distribution is determined by the capacitative network of the winding. For disk and helical windings, the capacitance to ground is usually much greater than the series capacitance through the winding. Under impulse conditions, most of the capacitive current flows through the capacitance to ground near the end of the winding, creating a large voltage drop across the line end portion of the coil.

The capacitance network for shell form and layer-wound core form results in a more uniform initial distribution because they use electrostatic shields on both terminals of the coil to increase the ratio between the series and to ground capacitances. Static shields are commonly used in disk windings to prevent excessive concentrations of voltages on the line-end turns by increasing the effective series capacitance within the coil, especially in the line end sections. Interleaving turns and introducing floating metal shields are two other techniques that are commonly used to increase the series capacitance of the coil.

Following the initial period, electrical oscillations occur within the windings. These oscillations impose greater stresses from the middle parts of the windings to ground for long-duration waves than for short-duration waves. Very fast impulses, such as steep chopped waves, impose the greatest stresses between turns and coil portions. Note that switching impulse transient voltages are two types—asperiodic and oscillatory. Unlike the asperiodic waves discussed earlier, the oscillatory waves can excite winding natural frequencies and produce stresses of concern in the internal winding insulation. Transformer windings that have low natural frequencies are the most vulnerable because internal damping is more effective at high frequencies.

Dielectric stresses existing within the insulation structure are determined using direct calculation (for basic geometries), analog modeling, or most recently, sophisticated finite-element computer programs.

Allowable stresses are determined from experience, model tests, or published data. For liquid-insulated transformers, insulation strength is greatly affected by contamination and moisture. The relatively porous and hygroscopic paper-based insulation must be carefully dried and vacuum impregnated with oil to remove moisture and gas to obtain the required high dielectric strength and to resist deterioration at operating temperatures. Gas pockets or bubbles in the insulation are particularly destructive to the insulation because the gas (usually air) not only has a low dielectric constant (about 1.0), which means that it will be stressed more highly than the other insulation, but also air has a low dielectric strength.

High-voltage dc stresses may be imposed on certain transformers used in terminal equipment for dc transmission lines. Direct-current voltage applied to a composite insulation structure divides between individual components in proportion to the resistivities of the material. In general the resistivity of an insulating material is not a constant but varies over a range of 100:1 or more, depending on temperature, dryness, contamination, and stress. Insulation design of high-voltage dc transformers in particular require extreme care.

10.1.6 Cooling

Removal of heat caused by losses is necessary to prevent excessive internal temperature that would shorten the life of the insulation. The following paragraphs cover the procedure for calculating the internal temperature of oil-insulated self-cooled power transformers of conventional core-type construction using radiators. Almost all modern power transformers have insulation systems designed for operation at 65°C average winding rise over ambient temperature and 80°C hottest-spot winding rise over ambient in an average ambient of 30°C. Older power transformers were designed for 55°C average winding rise/65°C hottest-spot winding rise over ambient.

The average temperature of a winding is the temperature determined by measuring the dc resistance of the winding and comparing it with the measurement previously obtained at a known temperature. The rise of the average temperature of a winding above ambient temperature is

$$U = B + E + N + T \tag{10-29}$$

where B = degrees Celsius rise of effective oil over ambient, E = degrees Celsius rise of average oil over effective oil, N = degrees Celsius rise of average coil surface over average oil, T = degrees Celsius rise of conductor over coil surface, and U = degrees Celsius rise of average conductor over ambient.

The effective oil temperature is the equivalent uniform temperature with equal ability to dissipate heat to the air. The effective oil temperature is approximately the average of the oil entering the top of the radiator and the oil leaving the bottom of the radiator. The oil temperature is approximately the same as the temperature of the adjacent radiator surface exposed to air. A smooth, vertical transformer-tank surface will dissipate heat to the air as follows:

$$D_B = 1.40 \times 10^{-3} B^{1.25} + 1.75 \times 10^{-3}\,(1 + 0.011A)B^{1.19} \tag{10-30}$$

where A = degrees Celsius ambient temperature, B = degrees Celsius effective oil rise over ambient, and D_B = watts per square inch dissipated to the air.

The first term of Eq. (10-30) covers heat transferred by convection. Usually, the radiator consists of parallel flattened tubes with limited accessibility to cooling air, and it is therefore necessary to multiply the first term by an experimentally determined friction factor (less than 1). The second term of Eq. (10-30) covers heat transferred by radiation, on the assumption of low temperature emissivity of 0.95, which applies to most painted surfaces commonly encountered. For any other value of low temperature emissivity this term should be multiplied by emissivity/0.95 (see Table 10-2). Usually, the radiator consists of parallel flattened tubes which radiate heat to each other. The net radiation of heat can be determined by considering the transformer and radiators replaced by a nonreentrant enveloping surface. If the second term of Eq. (10-30) is multiplied by the ratio of the area of the enveloping surface to actual surface (less than 1), the effect of reabsorption of radiation is eliminated. When radiation is small compared with convection, it can be assumed that A = 25°C and the $B^{1.19}$ can be replaced by $0.79B^{1.25}$, and Eq. (10-30) becomes

TABLE 10-2 Low Temperature Total Emissivity

Aluminum, highly polished	0.08
Copper	0.15
Cast iron	0.25
Aluminum paint	0.55
Oxidized copper	0.60
Oxidized steel	0.70
Bronze paint	0.80
Black gloss paint	0.90
White lacquer	0.95
White vitreous enamel	0.95
Green paint	0.95
Gray paint	0.95
Lampblack	0.95

$$B = \frac{100D_B^{0.8}}{(0.44F + 0.56\,V)^{0.8}} \quad °\text{C} \tag{10-31}$$

where V = ratio of envelope surface area to actual surface area and F = friction factor determined by experiment.

The temperature rise of average oil over effective oil, E, is usually negligible for normal transformer designs. It may become important if (1) the center of gravity of the radiators is not elevated sufficiently above the center of gravity of the core and coils, (2) there is unusual loss in the oil space over the core such as might result from high-current leads, (3) a winding has usually restricted oil ducts, or (4) pumps are used to circulate oil through the radiator without channeling the pumped oil through the oil ducts of the coil. For such cases, E is best evaluated by comparison with performance of previous designs.

The temperature rise of average coil surface over average oil, N, carries the loss in the coil through a film of stationary oil into moving oil. For a horizontal pancake coil (vertical axis), most of the heat escapes through the thin oil film on the upper surface and very little heat escapes from the lower surface. On the assumption that all the heat escapes from the upper surface, the temperature rise is

$$N = 13.2D_N^{0.8} \qquad °\text{C} \tag{10-32}$$

where D_N = watts per square inch dissipated from the coil to the oil.

For a vertical pancake coil (axis horizontal), the heat leaves both sides equally, and

$$N = 14D_N \qquad °\text{C} \tag{10-33}$$

The temperature rise of conductor over coil surface, T, carries the heat from the copper through the solid insulation applied to the conductor and the coil,

$$T = R_T t D_N \qquad °\text{C} \tag{10-34}$$

where D_N = watts per square inch dissipated from the coil to the oil, R_T = degrees Celsius per watt per inch thermal resistivity, and t = inch length of path.

The components of the winding rise over ambient are determined from Eqs. (10-31), (10-32), or (10-33), and (10-34) by using values of watts per square inch determined from the calculated losses and the design geometry. Then the total rise is determined from Eq. (10-29).

Oil circulation is as follows. The oil moves generally upward through ducts in the core and coils, rising in temperature as it goes. It moves generally downward through the radiators, falling in temperature as it goes (see Fig. 10-22). The space above the core and coils is filled with hot oil so that the height-temperature curve of the circulating oil forms a triangle *def*. The difference in weight of the two columns of oil which furnishes the circulating force is proportional to the area of the triangle

$$w = 2.50 \times 10^{-5}\, ml \tag{10-35}$$

where m = inches headroom, I = degrees Celsius top oil rise over average oil, and w = pound force per square inch circulating force.

I is established by the following relations:

$$L = 222IG_C \tag{10-36}$$

$$w = G_C R_H \tag{10-37}$$

$$I = 13.5\left(\frac{LR_H}{m}\right)^{1/2} \tag{10-38}$$

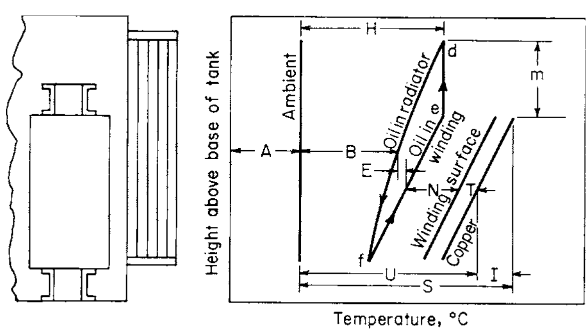


FIGURE 10-22 Oil-circulation diagram.

where G_C = gallons per minute rate of circulation of oil, L = watts loss, and R_H = friction opposing oil flow in pound-force per square inch per gallon per minute.

R_H is not easily evaluated except by test, but Eq. (10-38) is useful in evaluating the effect of changing L or m.

The limiting temperature rises are

$$H = B + E + I \tag{10-39}$$

$$S = B + E + N + T + I = U + I \tag{10-40}$$

where H = degrees Celsius top oil rise over ambient temperature and S = degrees Celsius hot-spot rise over ambient temperature.

Equation (10-40) gives the temperature of the top pancake coil. Values of N and T may need to be separately computed for this coil or for other coils if they have different loss density or different insulation than the main winding coils. If a coil other than the top coil is found to have higher conductor rise over adjacent oil, then appropriately reduced values of I should be used to calculate S for that coil.

Variation of temperature with load is as follows. If temperature rise conditions for rated load (or for any load) are known, temperature rises for any other load can be determined

$$B_2 = B_1\left(\frac{L_2}{L_1}\right)^{0.8} \tag{10-41}$$

$$N_2 = N_1\left(\frac{L_2}{L_1}\right)^{0.8} \tag{10-42}$$

$$T_2 = T_1\frac{L_2}{L_1} \tag{10-43}$$

$$I_2 = I_1\left(\frac{L_2}{L_1}\right)^{1/2} \tag{10-44}$$

where B_1, N_1, T_1, I_1, and L_1 correspond to the known condition and B_2, N_2, T_2, I_2, and L_2 correspond to the new condition.

The total loss should be used for L_1 and L_2 in Eqs. (10-41) and (10-44). The resistance and eddy loss only should be used for L_1 and L_2 in Eqs. (10-42) and (10-43). The exponent in Eq. (10-42) should be 1.0 for vertical pancake coils. At constant voltage the resistance loss varies with the square of the load kVA and with the resistivity as affected by average temperature of copper according to Eq. (10-25). The eddy loss varies with the square of the load and inversely with the resistivity of the copper according to Eq. (10-27). The stray loss varies with the square of the load and may be assumed to vary inversely with the resistivity of the copper, like the eddy loss. The core loss may be assumed unaffected by load or temperature. For many purposes, it is reasonable to assume that the entire load loss varies with the square of the load and with the resistivity.

Consider the following example. What is the temperature rise in a 30°C ambient at 80% load of a transformer with the following characteristics?

Load loss is 2 times no-load loss at 85°C

Load loss is assumed all resistance loss 85°C

Full-load temperature rises with 85°C losses are

$$B_1 = 45$$

$$N_1 = 15 \qquad U_1 = 65$$

$$T_1 = 5 \qquad S_1 = 80$$

$$I_1 = 15 \qquad H_1 = 60$$

On assuming (for a trial) that the average copper temperature will be 80°C, the relative load loss is

$$(0.8)^2 \times \frac{234.5 + 80}{234.5 + 85} = 0.630$$

and the relative total loss is

$$\frac{0.630 \times 2 + 1}{3} = 0.753$$

Then

$$B_2 = 45(0.753)^{0.8} = 35.9°\text{C}$$

$$N_2 = 15(0.630)^{0.8} = 10.4°\text{C}$$

$$T_2 = 5 \times 0.630 = 3.2°\text{C}$$

$$I_2 = 15(0.753)^{0.5} = 12.0°\text{C}$$

$$U_2 = 35.9 + 10.4 + 3.2 = 49.5°\text{C}$$

$$S_2 = 35.9 + 10.4 + 3.2 + 12.0 = 61.5°\text{C}$$

$$H_2 = 35.9 + 12.0 = 47.9°\text{C}$$

The average copper temperature is 49.5 + 30 = 79.5°C, which is close enough to the assumed 80°C.

Consider another example. What is the temperature rise in a 30°C ambient at 140% load for the same transformer? Upon assuming (for a trial) that the average copper temperature will be 140°C, the relative load loss is

$$(1.4)^2 \frac{234.5 + 140}{234.5 + 85} = 2.30$$

and the relative total loss is

$$\frac{2.30 \times 2 + 1}{3} = 1.87$$

Then

$$B_2 = 45(1.87)^{0.8} = 74.2°\text{C}$$

$$N_2 = 15(2.30)^{0.8} = 29.2°\text{C}$$

$$T_2 = 3 \times 2.30 = 6.9°\text{C}$$

$$I_2 = 15(1.87)^{0.5} = 20.5°\text{C}$$

$$U_2 = 74.2 + 29.2 + 6.9 = 110.3°\text{C}$$

$$S_2 = 74.2 + 29.2 + 6.9 + 20.5 = 130.8°\text{C}$$

$$H_2 = 74.2 + 20.5 = 94.7°\text{C}$$

The average copper temperature is 110.3 + 30 = 140.3, which is close enough to the assumed 140°C. These temperatures would be considered excessive for continuous loading and should not be continued for any extended time period.

Industry loading guides (ANSI/IEEE C57.92) provide tables that give winding hottest-spot temperature and top oil temperature for representative transformers over a wide range of ambient temperature and per unit loading conditions. These guides also contain a cautionary note that operation at hottest-spot temperatures above 140°C may cause gassing in the solid insulation and oil which could reduce the dielectric integrity of the transformer (Heinrichs 1979, McNutt, et al. 1980).

Transient thermal conditions must also be considered. Since transformer temperature takes many hours to stabilize after a change in loading, it is sometimes desirable to calculate the temperature during the transient period.

$$B_H = B_U - (B_U - B_0)\epsilon^{-h/hB} \tag{10-45}$$

$$h_B = \frac{K_B(B_U - B_0)}{L_D - L_0} \tag{10-46}$$

$$\begin{aligned} K_B &= 0.06H_{CC} + 0.04H_T + 1.33G_T \quad \text{for nondirected flow cooling} \\ K_B &= 0.06H_{CC} + 0.06H_T + 1.93G_T \quad \text{for directed flow cooling} \end{aligned} \tag{10-47}$$

where B_H = degrees Celsius effective oil rise after h hours, B_0 = degrees Celsius initial oil rise, B_U = degrees Celsius ultimate effective oil rise for the new load, G_T = gallons of oil, $\epsilon = 2.718$, h = hours after the change in load, h_B = hours time constant of the transformer, H_{CC} = pounds weight of core and coils, H_T = pounds weight of tank and fittings, K_B = watthours per degree Celsius thermal capacity of the transformer, L_D = watts dissipated at the new load, L_0 = watts loss at the initial condition ($h = 0$).

Equation (10-45) applies whether B_U is greater or less than B_0. However, if B_0 and L_0 are zero, then

$$B_H = B_U(1 - \epsilon^{-h/h_B}) \tag{10-48}$$

$$h_B = \frac{K_B B_U}{L_D} \tag{10-49}$$

The rise of average conductor over average oil ($N + T = Q$) is a single variable for transient calculations. The ultimate value is reached in 15 to 30 min

$$Q_H = Q_U - (Q_U - Q_0)\epsilon^{-h/h_Q} \tag{10-50}$$

$$h_Q = \frac{K_Q(Q_U - Q_0)}{L_0 - L_D} \tag{10-51}$$

$$K_Q = 0.05H_C\frac{a/2 + b}{b} \tag{10-52}$$

where a = square inches cross section of strand insulation, b = square inches cross section of copper strand, H_C = pounds weight of copper in winding, K_Q = watthours per degree Celsius thermal capacity of winding, L_D = watts dissipated at $h = 0$, L_0 = watts loss at $h = 0$, h = hours, h_Q = hours time constant of winding, Q_H = degrees Celsius average conductor rise over average oil after h hours, Q_0 = degrees Celsius initial average conductor rise over average oil, and Q_U = degrees Celsius ultimate average conductor rise over average oil.

Equation (10-50) applies whether Q_U is greater or less than Q_0. However, if Q_U and L_0 are zero, then

$$Q_H = Q_0\epsilon^{-h/h_B} \tag{10-53}$$

$$h_Q = \frac{K_Q Q_0}{L_D} \tag{10-54}$$

Now consider another example. If the transformer in the previous example operates in a 30°C ambient at 80% load until ultimate temperature conditions are established and then operates at 140% load, what will be the conditions after 4 h at 140% load? Assume nondirected flow cooling. It is now necessary to specify additional characteristics of the transformer as follows:

$$H_C = 20{,}000 \qquad a = 0.018$$

$$H_{CC} = 100{,}000 \qquad b = 0.090$$

$$H_T = 30{,}000$$

$$G_T = 10{,}000$$

$$\text{Total loss} = 150.000 \text{ W at } 85°\text{C}$$

Then $$K_B = 0.06 \times 100{,}000 + 0.04 \times 30{,}000 + 1.33 \times 10{,}000 = 20{,}500 \qquad (10\text{-}55)$$

Assume that the average winding temperature after 4 h is 125°C.

$$L_D = \frac{2 \times (1.4)^2(234.5 + 125)/319.5 + 1}{3} 150{,}000 = 270{,}540 \text{ W}$$

$$L_0 = 0.753 \times 150{,}000 = 112{,}950 \text{ W}$$

$$h_B = \frac{20{,}500\,(74.2 - 35.9)}{270{,}540 - 112{,}950} = 4.98 \text{ h}$$

$$B_H = 74.2 - (74.2 - 35.9)\epsilon^{-4/4.98} = 57.0°\text{C}$$

$$\left.\begin{aligned} N_H &= 29.2 \\ T_H &= 6.9 \\ I_H &= 20.5 \end{aligned}\right\} \text{(these reach ultimate value before 4 h)}$$

$$U_H = 57.0 + 29.2 + 6.9 = 93.1°\text{C}$$

$$S_H = 57.0 + 29.2 + 6.9 + 20.5 = 113.6°\text{C}$$

$$H_H = 57.0 + 20.5 = 76.5°\text{C}$$

Consider the following example of a load cycle. If the transformer examined in the preceding example operates in a 30°C ambient on a daily cycle of 20 h at 80% load and 4 h at 140% load, what are the temperature rises at the end of the 140% load period? Since the transformer time constant was approximately 5 h, it would be reasonable to assume that the oil temperature will have essentially stabilized after 20 h at 80% load. (The time constant for the transient from 140% load to 80% load will be essentially the same as the time constant for the transient from 80% load to 140% load.) As a result, the temperature rises at the end of the 140% load period during this cycle will be the same as those found in the previous example. Figure 10-23 shows the complete set of time-temperature curves for the 24-h cycle.

The short-circuit temperature rise is calculated on the assumption that all heat generated in the copper is stored in the copper until the short circuit is over

$$C_2 = 309.5 \times \left\{\left[\left(\frac{C_1 + 234.5}{309.5}\right)^2 + g\right]\epsilon^{M^2 s/10{,}800} - g\right\}^{1/2} - 234.5 \qquad (10\text{-}56)$$

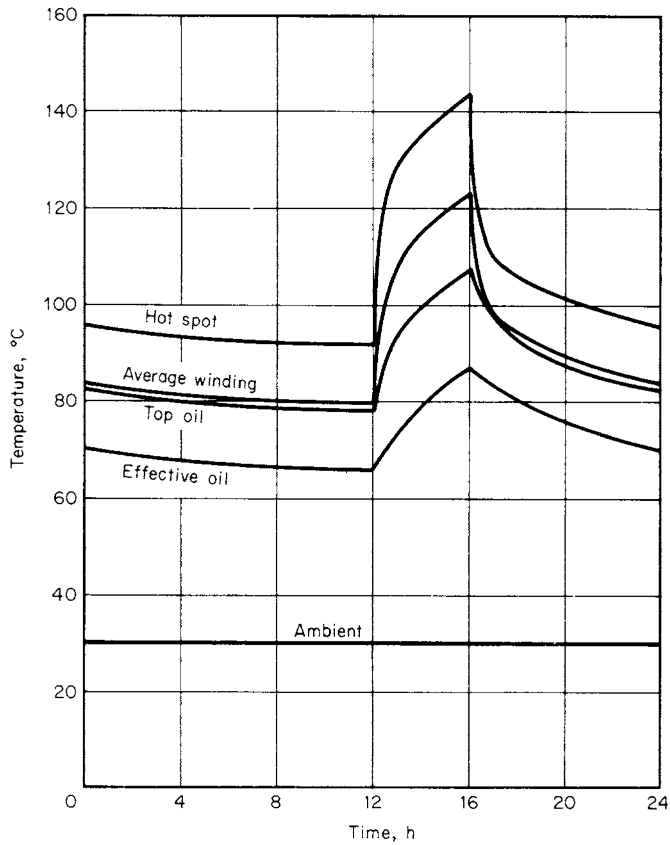


FIGURE 10-23 Transformer temperatures during the daily load cycle.

where, C_1 = degrees Celsius average temperature of winding at start of short circuit, C_2 = degrees Celsius average temperature of winding at end of short circuit, M = rms kiloamperes per square inch current density, s = seconds duration of short circuit, and g = ratio of eddy loss to resistance loss in winding at 75°C.

The temperature resulting from any short circuit may be read directly from Fig. 10-24, which is plotted from Eq. (10-56).

For example, if a short circuit resulting in 50 kA/in^2 current density is held for 3.5 s in a winding with 10% eddy loss ($g = 0.1$) at 75°C, and with a starting temperature of 75°C, what is the average temperature of the winding at the end of the short circuit?

On the curve for $g = 0.1$, at 75°C, the value of M^2s is 5300.

$$5300 + 50^2 \times 3.5 = 14{,}050$$

On the curve for $g = 0.1$, at $M^2s = 14{,}050$, the temperature is 246°C.

Fan-cooled transformers use external fans to improve heat dissipation from the radiators, and sometimes internal pumps to circulate the oil through the radiators (and sometimes also through cooling ducts in the core and coils). With fan cooling the effective oil-temperature rise B is determined from test data on the particular arrangement, instead of Eq. (10-31). With oil pumps I is determined from Eq. (10-38). It is usually possible to obtain 67% more capacity with fans and pumps running.

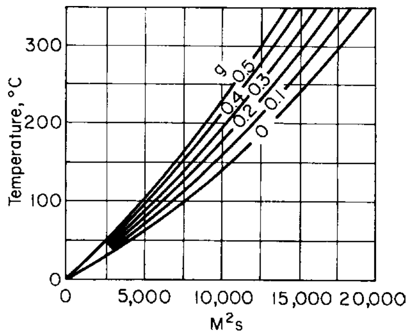


FIGURE 10-24 Temperature of windings after short circuit, with all heat stored, where g = ratio of eddy current loss to resistance loss at 75°C, M = rms kiloamperes per square inch current density, and s = seconds duration of short circuit. *Example*: For initial temperature of 75°C from the curve of $g = 0.1$, read $M^2s = 5300$. For $M = 50$ and $s = 3.5$; $(50)^2 \times 3.5 = 8750$. Total $M^2s = 14{,}050$. Then, from curve of $g = 0.1$, read final temperature: 246°C.

Forced-oil-cooled transformers use external oil-to-air heat exchangers requiring both air fans and oil pumps for all operating conditions. The effective oil rise B is calculated from the characteristics of the oil-to-air heat exchanger, and I is determined from Eq. (10-36). Forced-cooled transformers have no continuous load capacity without pump and fans.

Water-cooled transformers usually have the oil withdrawn from the transformers at the top of the tank, pumped through an external cooler, and returned to the bottom of the tank. The temperature drop of the oil in passing through the cooler is

$$Y = \frac{L}{G_c \times 111} \tag{10-57}$$

where Y = degrees Celsius drop of oil in cooler, L = watts loss, and G_c = gallons-per-minute rate of circulation of oil.

The effective oil-temperature rise B may be calculated from the characteristics of the oil-to-water heat exchanger. The top oil rise over average oil, I, is $Y/2$. The other components of temperature rise are calculated as for a self-cooled transformer.

Operation at high altitude increases the effective oil rise of air-cooled transformers. ANSI C57 provides for a compensating correction of 0.4% of rated kVA for self-cooled transformers or 0.5% of rated kVA for forced-air-cooled or forced-oil-cooled transformers for each 330 ft of additional altitude above 3300 ft altitude.

Effect of Tank Color. Most paint used on transformers has a low temperature emissivity of about 0.95. Metallic surfaces, particularly polished surfaces, have less low temperature emissivity and will cause correspondingly higher oil-temperature rise. The same is true of aluminum or bronze paint. For these cases, Eq. (10-31) can be used with V' substituted for V, where V' is $V/0.95$ multiplied by the low temperature emissivity from Table 10-2. On large power-transformers with many radiators or heat exchangers the effect is small.

For transformers exposed to intense sunlight, the additional temperature rise resulting from the use of aluminum paint is largely offset by the fact that aluminum paint absorbs only about 55% of the impinging solar radiation, whereas most commonly used paints absorb about 95%.

10.1.7 Load-Tap Changing

Ratio Changes with Shifted Taps. In transformers designed for maintaining a constant voltage on a power system, the ratio of transformation is usually changed by increasing or decreasing the number of active turns in one winding with respect to another winding. Since the turn ratio of the transformer must be changed without interfering with the load, means are provided for shunting the load current from one winding tap to the next. For this purpose, an auxiliary *preventive autotransformer* is generally used, designed to limit the resulting circulating current to a safe value during the interval when two adjacent taps are bridged. Because of the circulating and the load current which passes through the current-limiting impedance, arcing always takes place as the power circuit is transferred from tap to tap.

Although a variety of switching equipment and transformer connections have been used for the purpose of changing taps under load, the underlying principle remains unchanged and is shown by the transformer connection in Fig. 10-25.

Example. To move from transformer tap A to B, it is first necessary to close the circuit to B, as shown in Fig. 10-25, before opening the circuit at A. During the interval when A and B are both closed on adjacent taps, a circulating current flows through and is limited by the impedance on the loop composed of the tap winding *AB* and autotransformer *C*. With both ends of the autotransformer connected to *A*, the load current divides equally between the two halves of the autotransformer. Since the current flows in opposite directions, a negligible amount of reactance is introduced into the circuit and the only loss is the I^2R due to the 50% load current in each half of the autotransformer winding. With *A* closed and *B* open, all the load current flows through one-half of the autotransformer, magnetizing the autotransformer and thereby introducing into the circuit the induced voltage. It is important, therefore, that the magnetizing reactance be kept as low as possible to avoid excessive arcing duty on the circuit-interrupting device.

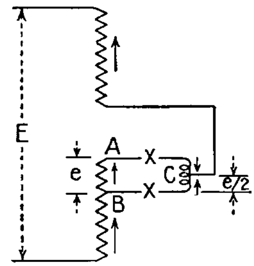


FIGURE 10-25 Bridging position for ratio change under load.

With *A* and *B* closed on adjacent taps, the tap voltage *e* is impressed on the autotransformer *C* and causes a circulating current to flow through the impedance loop. Because of the autotransformer action, a voltage midway between *A* and *B* is impressed in the circuit. The load current again divides equally through the autotransformer windings.

To avoid an excessive voltage drop through the autotransformer when one side is open and at the same time to keep the circulating current at a low level when in the bridging position, the autotransformer is usually designed with an air gap in the magnetic circuit to get a magnetizing current of approximately 60% of the normal full-load current.

Voltage across Autotransformers. Figure 10-26 shows the voltage relations across an autotransformer and switching contacts during a tap-changing cycle using an autotransformer designed for 60% circulating current and with 100% load current at 80% power factor flowing through it. Perfect interlacing between the autotransformer halves is assumed, and the voltage drop due to resistance of the autotransformer winding is neglected.

A study of Fig. 10-26 will disclose the fact that increasing the magnetizing reactance of the autotransformer to reduce the circulating current will

1. Increase the voltage across the full autotransformer winding
2. Increase the voltage to be ruptured
3. Introduce undue voltage fluctuations in the line

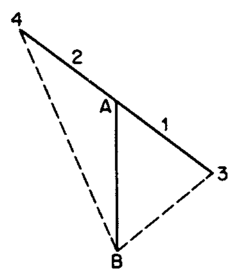


FIGURE 10-26 Vector relations for bridging position *AB*—voltage across adjacent taps; *A*-1 and *A*-2— reactance volts due to load current in only half the autotransformer winding; *A*-3 and *A*-4—induced voltage across full auto transformer winding; *B*-4—voltage ruptured when bridging position is ruptured at *A* (Fig. 10-25); *B*-3—voltage ruptured when bridging position is ruptured at *B* (Fig.10-24).

Since *B*-4 and *B*-3 represent the voltages appearing across the arcing contacts when the bridging position is opened at *A* and *B*, the voltage-rupturing duty will increase with

1. Increase in voltage between adjacent taps
2. Increase in load
3. Decrease in power factor of the load
4. Decrease in the magnetizing current for which the autotransformer is designed

Load-Tap-Changer Motor Mechanisms. The mechanism which drives the tap changer, and the control of this mechanism, must be designed so that a tap change, once initiated, is certain to be completed.

The mechanical coupling between the operating motor and the tap-changing switches may be through fixed-ratio gears, Geneva gears, cams, springs, or combinations of these. All mechanisms require means for keeping the motor energized until the change of tap is accomplished and for bringing the tap changer to rest on each operating position. The degree of permissible coasting of the motor is determined by the motor mechanism and the switch design.

The need for extremely accurate stopping of the motor is avoided by arranging the parts so that the motor may coast somewhat after the operating position is reached without moving the tap changer around the operating position. This may be accomplished by the inactive sectors of Geneva gears or cams or by the motor travel inherently involved in recharging a spring. Motor mechanisms are provided with limit switches and mechanical stops to prevent operation beyond the limit positions. Operation counters and position indicators are standard auxiliaries on most tap chargers. On large station-type units where the control devices are generally mounted on a remote-control panel, remote position indicators, either of the self-synchronizing type or of the digital type are generally used.

Automatic Control for Tap Changers. It is usual practice to use some sort of voltage measuring device to control the operation of the motor which drives the tap changer. Such devices may be mechanical, balancing the force of a solenoid actuated by the voltage against weights or springs, or they may be an electrical network, usually a bridge circuit which balances against the voltage of a Zener dioide. With either type of device, a voltage higher than a desired upper limit will start the tap-changer driving motor to change to the next lower tap voltage; similarly, a voltage lower than the desired lower limit will cause a change to the next higher tap.

The circuit usually includes a time delay to prevent tap changes, which would occur unnecessarily during very short time variations in voltage. It also may include a line drop compensator to facilitate maintaining the voltage within a given band at a point (load center) some distance from the transformer. The line-drop compensator introduces a signal into the voltage regulating relay circuitry. This represents the voltage drop due to line impedance between the transformer and the load center.

The voltage-regulating relay (or contact-making voltmeter) should be adjusted so that the voltage bandwidth, or spread between voltages at which the raising and lowering contacts close, will be not less than the percentage transformer tap plus an allowance for irregular voltage variations. For example, a tap-changing transformer with $1\frac{1}{4}\%$ taps should have a minumum voltage bandwidth of approximately $1\frac{1}{4}\% + \frac{1}{2}\% = 1\frac{3}{4}\%$.

In addition, the voltage-regulating relay may contain a component for use when load tap-changing transformers are operated in parallel. In this case, the tap changers must be controlled so that they are approximately on the same tap position. The component, a paralleling reactor, is used with external circuitry to detect, and generate a signal to minimize, circulating current that results when the tap changers are not on like positions.

Voltage Control, a Part of the Power Transformer. The simplest and generally the least expensive connection for voltage control is to provide the necessary taps in the power transformer. For single- or 3-phase delta connection the taps are preferably located on the interior of the winding (Fig. 10-27) so as to avoid the abnormal voltage stresses to which end coils are usually subjected. In the Y connection the taps may be placed at the neutral end of the winding, and if the neutral is to be solidly grounded, it becomes possible, by locating the taps next to ground, to use load changers designed with greatly reduced insulation; thus, for example, 15-kV apparatus may be placed in the grounded neutral end of a much higher-voltage circuit.

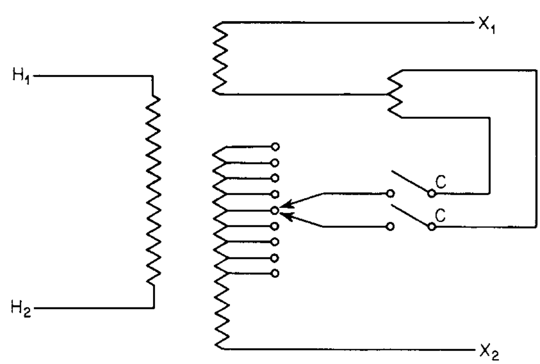


FIGURE 10-27 Tap-changing equipment in the middle of the winding.

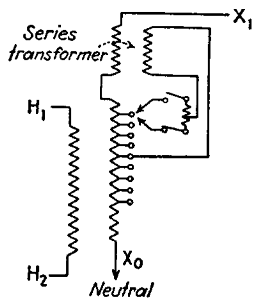


FIGURE 10-28 Tap-changing circuit with taps located in the interior of the transformer winding and an auxiliary series transformer to bring the current and voltage duty on equipment within rating limits.

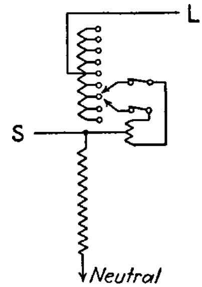


FIGURE 10-29 Regulating single-core autotransformer with taps located in the series winding and circuit connected for boost and buck.

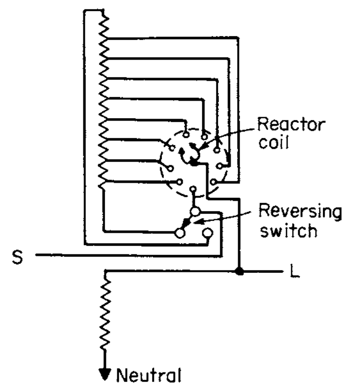


FIGURE 10-30 Regulating single-core autotransformer with a reversing switch to obtain buck and boost.

If the rated current of the transformer exceeds that of the switching equipment, a series transformer may be used (Fig. 10-28). Excitation is derived from taps inserted in the secondary of the power transformers, and by means of the series transformer, the desired voltage is inserted into the circuit. Thus, if a ratio of 3:1 exists in the series transformer, the current handled by the switching equipment becomes one-third of the current in the line.

Regulating Transformers (Single-Core). When a power transformer is not available or it is not desirable to equip the power unit with voltage control, regulating autotransformers are used. In the simplest of these, the necessary taps and switches are placed in the series winding of an autotransformer (Fig. 10-29). For 3-phase circuits, in order that the derived voltage may be in phase with circuit voltage, a Y connection is commonly used, and hence all the precautions necessary to safeguard the operation of Y-connected autotransformers should be observed. A tertiary winding may or may not be provided, depending on circuit conditions. As the series winding is inserted in the line, adequate insulation must be provided for the tap-changing equipment and taps against the abnormal voltages to which the circuit is subjected.

Economy in transformer size may be obtained by means of a reversing switch which functions to reverse the connections to the series winding when the regulator is passing through the neutral position. The circuit is so designed and the mechanical sequence is such that the reversing switch operates without rupturing current. The connection diagram (Fig. 10-30) shows the load tap changer provided with nine taps, which gives 17 full-cycle or 33 half-cycle positions. The ratio adjuster is designed with contacts uniformly spaced on the circumference of the circle so as to permit motion through two revolutions.

The Series Transformer. In many instances the voltage of the circuit is greater than that for which the switching equipment is designed, and in others the current to be handled exceeds the safe limits of operation. In either case, voltage control can be obtained without the design of special switching equipment by using an insulating series transformer (Fig. 10-31) in addition to the exciting transformer, the combination functioning, as far as the circuit is concerned, like an autotransformer. The primary of the exciting transformer is generally connected in Y in order that the derived voltages may be in phase with circuit voltages. The secondary of the exciting transformer provided with the regulating taps is usually connected in delta. The local circuit, consisting of the secondary of the exciting transformer with its taps and the primary of the series transformer being insulated from the main circuit, may be designed for the voltage and current best suited for the available switching equipment. Because of the additional cost and losses of the series transformer, it is used only when the voltage

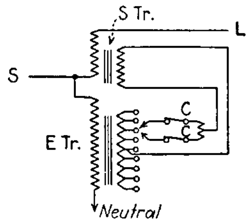


FIGURE 10-31 Exciting transformer with taps in the secondary and a series transformer forming complete isolation for tap-changing equipment.

or current limitations of the switching equipment demand it and when the control cannot be inserted in the grounded neutral of the transformer bank.

Tap-Changer Designs for Moderate kVA and Current. In the smaller ratings, where both the voltage and the current are moderate, the energy to be ruptured in switching from tap to tap becomes relatively so small that light and simple equipments are feasible. A variety of mechanical designs, together with special circuits, has been evolved with the purpose of providing simpler, smaller, and inherently less expensive equipments. The following may be noted:

1. Designing the tap changer so that it is capable of rupturing the current directly on the same switches which select the taps
2. Designing the circuit so that the tapped winding is reversed in going from maximum to minimum range, thereby securing a substantial reduction in the rating of core and coils for a given output
3. Using higher switching speed, by means of which the life of the arcing contacts is increased

Tap Changers Designed to Interrupt Current. The contactors *C* (Fig. 10-27) operate to open the switching circuits so that there is no interrupting duty on the selector contacts which connect to the transformer taps. When the rated current is moderate, it becomes possible to rupture the current directly on the tap-selector switches and thus obtain a major economy in the cost of the mechanical equipment. This method is shown in Fig. 10-30.

High-Speed Switching. Large units include contactors with high-speed contacts which serve as extinction devices that are specially designed to repeatedly interrupt the high currents and voltages encountered. They may be single- or multiple-break contactors operating in oil or in air with magnetic-arc chutes, or oil-blast contactors. Vacuum switches with their longer life (reduced maintenance) have become widely used. In small units, however, the arcing duty is mild. It is nevertheless necessary to keep in mind that mild arcing duty in the smaller equipment is partly offset by the likelihood of greater frequency of operation. Such units are usually equipped with full automatic control; they are likely to be located on distribution circuits where the voltage is more erratic. Many of them are located on the lines at considerable distances from substations, and some of them are placed on poles. It is desirable, therefore, to reduce maintenance to a minimum. For these reasons, it is necessary to provide means for high-speed switching on the smaller units where the tap-selector switches are used to rupture current. High-speed action of the tap-changer switches is obtained through Geneva gears or cams which bring the contact fingers to the required high speed at the moment of parting or through a spring drive in which the motor is used to store energy with the release of a spring snapping the contact fingers from one shelf to the next. By these means, the duration of the switching arc may be reduced to one or two contactors correspondingly reducing the amount of contact burning and increasing the life of the contacts.

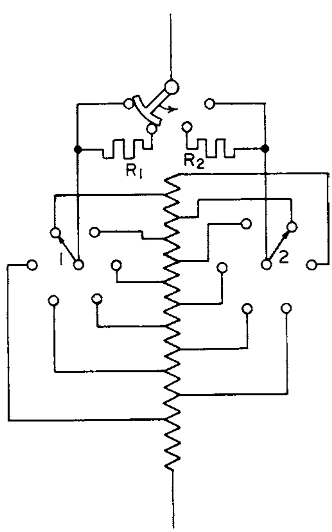


FIGURE 10-32 Tap-changing circuit employing tap selectors and contact, or diverter switch, and resistors to bridge the taps during switchover.

Use of Resistors. Another method, used more frequently in load tap changers of European design, makes use of resistors to bridge the tap instead of a preventive autotransformer. Figure 10-32 shows the tap selectors 1 and 2 connected to alternate taps in the winding. The contactor, or diverter switch, shown connected to 1 and R_1, progressively connects only to R_1, then to R_1, and R_2, then to R_2, and finally to R_2 and tap selector 2. For the next tap change, tap selector 1 moves to an adjacent tap and is then followed by the diverter switch operating in a reverse manner from R_2 and 2 back to R_1 and 1. Since the resistors are designed to carry current for only a very short time, the diverter switch is usually spring-actuated and moves through its sequence in a few cycles of 60-Hz current. This method has the advantage of relatively small-size resistors but requires a transformer tap for each operating voltage, while the autotransformer circuit uses the tap bridging position for an operating voltage and thus requires half the number of transformer taps.

Applications for Voltage Control and Equipment. The control of transformer ratio under load is a desirable means of regulating the voltage of high-voltage feeders and of primary networks. It may be used for the control of the bus voltage in large distributing substations. It finds a wide field of application in controlling the ratio on step-up transformers operating from power stations whose bus voltage must be varied to suit local distribution.

In industrial work, it is used for the control of current in a variety of furnace operations and electrolytic processes. It also furnishes a convenient means for voltage regulation of concentrated industrial loads.

A lot of load tap-changer equipment is installed at points of interconnection between systems or between power stations, in order to control the interchange of reactive current, or, in other words, to control the power factor in the tie line. This reactive current may be highly undesirable, especially as it may add to the burden on a fully loaded generating system. It can be increased, eliminated, or reversed by inserting a suitable small ratio of transformation between the systems. It can be varied in amount and in direction of flow to suit varying system conditions, if this ratio is variable and under the control of a station operator. Inserting such a ratio of transformation in a tie line by means of tap-changing equipment is equivalent in its effect on the flow of reactive current to raising or lowering the voltage on one of the systems. Current can be exchanged at any power factor from zero lag to zero lead, without interfering with the voltage maintained on either system.

Transformers for Phase-Angle Control. Tap-changing equipment is sometimes used in a loop system, for phase-angle control, for the purpose of obtaining minimum losses in the loop due to unequal impedances in the various portions of the circuit.

Transformers used to derive phase-angle control do not differ materially, either mechanically or electrically, from those used for inphase control. In general, phase-angle control is obtained by interconnecting the phases, that is, by deriving a voltage from one phase and inserting it in another.

The simple arrangement given in Fig. 10-33*a* illustrates a single-core delta-connected autotransformer in which the series windings are so interconnected as to introduce into the line a quadrature voltage. One phase only is printed in solid lines so as to show more clearly how the quadrature voltage is obtained. The terminals of the common winding are connected to the midpoints of the series winding in order that the inphase voltage ratio between the primary lines *ABC* and secondary lines *XYZ* is unity for all values of phase angle introduced between them.

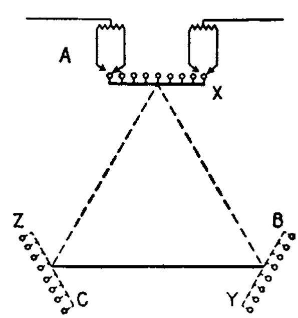


FIGURE 10-33*a* Phase-shifting regulating transformers; single-core delta-connected common winding for low-voltage systems.

As large high-voltage systems have become extensively interconnected, a need has developed to control the transfer of real power between systems by means of phase-angle-regulating transformers. The most commonly used circuit for this purpose is the two-core, four-winding arrangement shown in Fig. 10-33*b*. The high-voltage common winding is Y-connected, with reduced insulation at the neutral for economy of design, and a series transformer is employed so that low-voltage-switching equipment may be used.

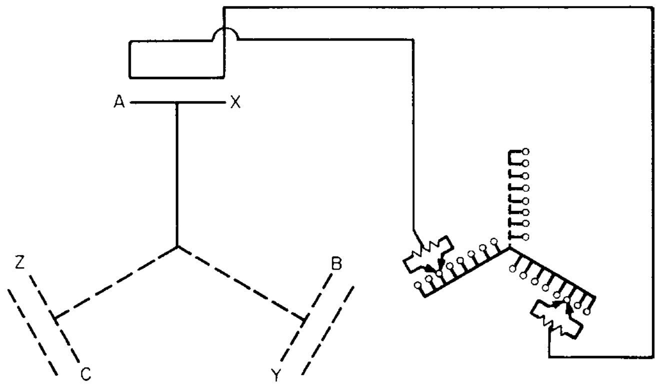


FIGURE 10-33*b* Phase-shifting regulating transformer; two-core Y-connected common winding for high-voltage systems.

10.1.8 Audible Sound

Source of Sound. Transformers, although they are classed as static apparatus, vibrate and radiate audible sound energy. In general, there are three different sources of transformer noise:

1. *Core vibration due to magnetostriction*. Core steel laminations undergo elongations and contractions (magnetostriction) as flux through them varies. This magnetostriction is nonlinear and independent of flux direction. Hence, noise is emitted in even multiples of the excitation frequency, that is, 120, 240, 360 Hz, etc., for a 60-Hz power system. The harmonic components decrease in magnitude as the mode of vibration goes up. However, an overexcited transformer or core-resonance may produce abnormally high third or higher harmonic frequencies.
2. *Noises from cooling equipment*. All rotary machinery on a transformer, including fans and pumps, produces noise with a broadband frequency spectrum. This "white noise" can have various magnitude and directionality depending on the design of the fans and pumps and on their arrangement.
3. *Coil vibration from energization*. Coils in a transformer are under cyclical stresses due to stray fluxes. The resultant motion resembles a vibrating spring and can also emit noise with harmonics of 120 Hz. However, this component is generally much lower than the previous two sources unless the transformer has a low induction level and high power ratings.

Sound Measurement. Sound waves produce small fluctuations in the atmospheric pressure which are sensed by the human ear.

Sound-level-measuring equipment as specified by ANSI Standard S1.4 consists of a microphone, amplifier, frequency weighting network, and indicating meter.

The most common type is A-weighing (Fig. 10-33*c*). It represents the sensitivity of the young adult ears to moderate sound levels over most of the audible spectrum. A linear response gives the actual sound intensity level. One-third octave and narrowband sound-level measurements are used to identify the source of an unexpectedly noisy transformer.

Standard Transformer Sound Level. ANSI/IEEE C57.12.90 specifies the method for measuring the average sound level of a transformer. The measured sound level is the arithmetic average of a number of readings taken around the periphery of the unit. For transformers with a tank height of less than 8 ft, measurements are taken at one-half tank height. For taller transformers, measurements are taken at one-third and two-thirds tank height. Readings are taken at 3-ft intervals around the string periphery of the transformer, with the microphone located 1 ft from the string periphery and

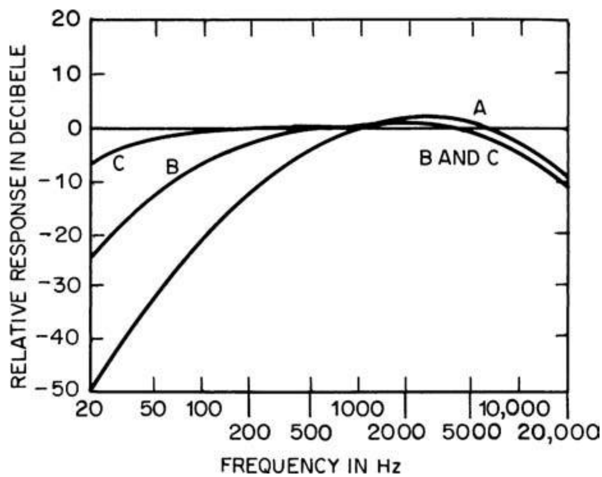


FIGURE 10-33*c* Weighting curves; A-weighting reduces the intensity of noise toward the lower end of the audible spectrum.

6 ft from fan-cooled surfaces. The ambient must be at least 5 and preferably 10 dB below that of the unit being measured. There should be no acoustically reflecting surface, other than ground, within 10 ft of the transformer. The A weighting network is used for all standard transformer measurements regardless of sound level.

NEMA Publication TR 1 contains tables of standard sound levels. For oil-filled transformers, from, 1000 to 100,000 kVA, self-cooled (400,000 kVA, forced-oil-cooled) standard levels are given approximately by Eq. (10-58)

$$L = 10 \log E + K \tag{10-58}$$

where E = equivalent two-winding, self-cooled kVA (for forced-oil-forced-air-cooled units, use 0.6 × kVA), K = constant, from Table 10-3, and L = decibel sound level.

Example. A transformer rated 50,000 kVA self-cooled, 66,667 kVA forced-air-cooled, 83,333 kVA forced-oil-forced-air-cooled, at 825 kV BIL, would have standard sound levels of 78, 80, and 81 dB on its respective ratings.

Public Response to Transformer Sound. The basic objective of a transformer noise specification is to avoid annoyance. In a particular application, the NEMA Standard level may or may not be suitable, but in order to determine whether it is, some criteria must be available. One such criterion

TABLE 10-3 Values of K for Eq. (10-58)

High-voltage winding BIL, kV	Self-cooled and water-cooled ratings	Forced-air and forced-oil-forced-air-cooled 25% to 35% above self-cooled rating	Forced-air and forced-oil-forced-air-cooled 67% above self-cooled rating or without self-cooled rating
350 and below	28	30	31
450 to 650	30	32	33
750 to 825	31	33	34
900 to 1050	32	34	35
1175	33	35	36
1300 and above	34	36	37

is that of audibility in the presence of background noise. A sound which is just barely audible should cause no complaint.

Studies of the human ear indicate that it behaves like a narrowband analyzer, comparing the energy of a single frequency tone with the total energy of the ambient sound in a critical band of frequencies centered on that of the pure tone. If the energy in the single-frequency tone does not exceed the energy in the critical band of the ambient sound, it will not be significantly audible. This requirement should be considered separately for each of the frequencies generated by the transformer core.

The width of the ear-critical band is about 40 Hz for the principal transformer harmonics. The ambient sound energy in this band is 40 times the energy in a 1-Hz-wide band. The sound level for a 1-Hz bandwidth is known as the "spectrum level" and is used as a reference. The sound level of the 40-Hz band is 16 dB (10 log 40) greater than the sound level of the 1-Hz band. Thus, a pure tone must be raised 16 dB above the ambient spectrum level to be barely audible.

The transformer sound should be measured at the standard NEMA positions with a narrow-band analyzer. If only the 120- and 240-Hz components are significant, an octave-band analyzer can be used, since the 75- to 150-Hz and 150- to 300-Hz octave bands each contain only one transformer frequency. The attenuation to the position of the observer can be determined.

The ambient sound should be measured at the observer's position. For each transformer frequency component, the ambient spectrum level should be determined. An octave-band reading of ambient sound can be converted to spectrum level by the equation

$$S = B - 10 \log C \tag{10-59}$$

where B = decibels octave-band reading, C = hertz octave bandwidth, and S = decibels spectrum level.

Example. Consider the following case:

Transformer sound at 120 Hz by NEMA method = 72 dB

Transformer-sound attenuation to observer = 35 dB

Ambient sound at the 75- to 150-Hz octave band = 36 dB

72 − 35 = 37 dB at the observer's position

36 − 10 log (150 − 75) = 17.3-dB ambient spectrum level

The 120-Hz transformer sound at the observer's position exceeds the ambient spectrum level by 19.7 dB. This is 3.7 dB greater than the 16-dB differential which would result in bare audibility; thus the transformer sound will be audible to the observer.

When transformer sound exceeds the limits of bare audibility, public response is not necessarily strongly negative. Some attempts have been made to categorize public response on a quantitative basis when the sound is clearly audible (Schultz and Ringlee 1960). For a case where specific knowledge of transformer- and ambient-sound-level frequency composition is not available, some more general guidelines are useful. Typical average nighttime ambient-sound levels for certain types of communities have been established. These are 30 dB for a "quiet suburban," 35 dB for a "residential suburban," and 40 dB for a "residential urban" community. All sound levels are based on the A scale of weighing. Calculations for typical transformer frequency distributions have been made to determine the nighttime transformer noise which will be audible 50% of the time in these communities. The results are 24 dB for quiet suburban, 29 dB for residential suburban, and 34 dB for residential urban. The NEMA standard sound level can be corrected for attenuation with distance to the nearest observer and checked against the above guides for audibility.

The broadband sound from fans, pumps, and coolers has the same character as ambient sound and tends to blend in with the ambient. While the noise from cooling equipment may be audible to a neighboring observer, it will seldom, if ever, cause a complaint.

Sound Attenuation with Distance. A point source in a free field radiates sound in spherical waves. The resultant sound pressure varies inversely with the square of the distance from the source; thus the sound level is reduced by 6 dB for each doubling of distance. The sound of auxiliary cooling equipment follows this relation for decrement with distance, since it is the sum of point-source sound contributions.

The transformer tank, which radiates vibrational energy from the core, is a more complex sound source and does not appear as a point source except at substantial distance from the tank. The modes of tank vibration are complicated, and various parts of the tank may act as independent sources, with different amplitudes, phase relations, and frequencies. Studies of scale models (Johnson et al. 1956) and full-size units have uncovered certain useful relationships as follows:

$$A = 20 \log \frac{2.83\,D}{Q} \tag{10-60}$$

$$Q = 1.7(WH)^{1/2} \tag{10-61}$$

where A = decibels attenuation for distance exceeding Q, D = distance from transformer to observer, H = height of transformer tank, Q = critical distance from transformer beyond which it appears as a point source, and W = width of transformer tank perpendicular to a line from transformer to observer.

Equations (10-60) and (10-61) apply in the absence of wind, temperature gradients, and reflecting surfaces other than ground. Each of these factors may significantly influence the observed sound level at a distance from the source, but not always in predictable fashion.

Site Selection. There are a number of methods available for avoiding transformer-noise complaints. Some of the discussion in the previous paragraphs suggests that potential noise problems should be considered when the substation site is selected. It may be possible to take advantage of attenuation with distance to reduce the transformer sound at the nearest observer position to an inaudible level. It may also be possible to choose the site in a location where the normal ambient noise will mask the transformer sound. If these possibilities are kept in mind during the planning stages, more expensive solutions to noise problems may be avoided later.

Design Measures. Manufacturers have at their disposal a variety of means of obtaining sound reduction. Most measures aim at reducing noise generation.

1. *Reducing core vibration*. Since magnetostriction is a function of flux intensity, a manufacturer's first option is to reduce induction levels of transformers. This has an additional advantage of reducing no-load losses. Alternatively, grades of steel having a different magnetostrictive characteristics can be substituted for the same induction-level design. A step-lap design can also reduce noise emission from joints. Finally, the designer has to anticipate the natural frequencies of the core mechanical structure and avoid their coincidence with harmonics of 120 Hz.
2. *Reducing cooling equipment noises*. The most significant noise reduction measure for cooling equipment is to reduce fan rotational speed or adjust the fan blade incidence angle. There is ample supply of low-noise designs, ranging from low speed to encased fans, from which manufacturers can choose.

When all possibilities of noise emission are exhausted and still further noise reduction is required, some sort of a mass-damper or absorption system has to be incorporated on or outside the tank structure. Moderate reductions can be realized by the use of barriers within the tank. Some of these are "soft" barriers, which operate on the principle of absorbing vibrational energy from the core and reducing its transmission to the tank. Others are "mass" barriers, which operate on the principle of loading the tank to decrease its magnitude of vibration for given energy transmission from the core. To achieve large sound reductions (as much as 25 to 30 dB), some manufacturers employ complete external enclosures of steel. For smaller substation units, these enclosures can be preassembled and shipped in place over the transformer tank.

Improving Existing Installation. To reduce the sound level of an existing transformer, the most satisfactory method has been found to be the erection of barrier walls on one or more sides of the transformer. The attenuation which can be achieved depends on the transmission loss through the barrier, the diffraction over and around the barrier, and the pressure buildup between the tank and the barrier.

Transmission loss through a barrier wall is a function of the mass of the wall. Structural requirements of most practical masonry barriers ensure sufficient mass to produce 25 to 40 dB attenuation through the wall. The effectiveness is usually limited by diffraction around the edges of the barrier. A theoretical method for calculation of attenuation as limited by diffraction has been formulated as follows:

$$N = \frac{2}{\lambda}[(M^2 - U^2)^{1/2} - M + (G^2 - U^2)^{1/2} - G] \tag{10-62}$$

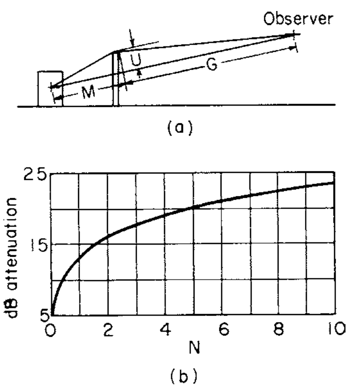


FIGURE 10-34 Effectiveness of a barrier in reducing noise level: (*a*) identification of dimension for calculation of the dimensionless parameter *N* from Eq. (10-62); (*b*) determinations of attenuation in decibels.

where *M*, *U*, and *G* are as defined in Fig. 10-34, in any convenient unit (feet or inches), λ = wavelength of harmonic under investigation, in units consistent with *M*, *U*, and *G*, and *N* = dimensionless parameter given in Fig. 10-34.

The calculation procedure is to determine *N* from the equation and then find the corresponding attenuation from Fig. 10-34*b*.

Test results on models and full-size transformers with two- and three-wall barriers correlate reasonably with Eq. (10-62). Data on four-wall enclosures generally do not correlate. It has been found that approximately 10-dB attenuation can be achieved with a four-wall enclosure having walls 5 ft higher than the transformer.

Enclosures with fewer than four walls should extend at least a distance *M* beyond the tank, so that attenuation will be limited by diffraction over the top rather than around the ends of the barrier. It should be noted that the sound level on the open side of this type of enclosure will be increased above what it was without the enclosure. Energy is redirected from the critical side of the transformer to the less critical side.

The effective attenuation of an enclosure can be reduced by pressure buildup between the tank and the barrier. The buildup is the result of reflection from hard wall surfaces and reinforcement of direct and reflected waves. Buildup will be most pronounced for spacings between tank and barrier walls which are multiples of the half wavelength of any of the principal sound frequencies. Such spacings should be avoided. Sound-absorbent lining on the interior surface of the barrier walls is helpful in reducing or eliminating buildup.

Masonry enclosures can also be used to hide substation transformers and associated equipment and in that way alleviate complaints which are based on appearance in addition to noise. Some utilities use a three-sided enclosure which resembles the houses in the neighborhood (Buck 1959). A casual observer on the street may not detect the presence of the substation.

10.1.9 Partial Discharges

Partial discharges may take place in liquid or gaseous dielectrics when the dielectric stress at the point of maximum stress concentration reaches the breakdown level but when complete breakdown of the dielectric is prevented because the dielectric stress decreases very rapidly away from the point of maximum stress concentration or because a solid dielectric intervenes. One form of such partial discharge, in air around a small conductor at high voltage, has been called “corona” because of its appearance as a visible glow around the conductor surface.

The local breakdown in the region of stress concentration ionizes a path (forms a streamer) in a very short time (microseconds), effectively short-circuiting a small region of the dielectric, and a pulse of current appears in the main dielectric circuit, reflecting the instantaneous short-circuiting of part of the circuit capacitance.

Partial discharges usually are accompanied by chemical decomposition of the liquid or gas, and sometimes they cause erosion of the adjacent solid insulation. A partial discharge in oil usually causes chemical breakdown, with the formation of carbon and gas, and unless the gas is immediately able to escape, more severe discharges in the gas itself may lead to complete breakdown of the insulation structure.

The presence of gas bubbles in the insulation of an oil-insulated transformer may result in partial discharges; this is the reason for particular attention to filling transformers with oil under vacuum.

Partial discharges may also be caused by wet fibers or any small conducting particles which distort the electric field and cause local points of stress concentration.

Partial discharges can be detected when they occur within the insulation of a transformer by any of a number of schemes which detect or measure either the pulse of current or the momentary loss of voltage at the transformer terminal. The charge transfer at a terminal can be measured in picofarads, but this generally does not give the actual transfer of charge which occurs somewhere within the transformer. Two techniques are commonly used to measure partial discharge activity. NEMA Publication 107 describes a method for measuring the equivalent high-frequency voltage, usually at 1 MHz, which appears at the terminals of the transformer, while ANSI/IEEE CS7.113 presents a trial-use guide to measure picocoulombs (apparant charge). The apparent charge technique is more sensitive to partial discharges occurring within the winding but also can be more susceptible to external signals. For both techniques, for power transformers the coupling capacitor can be replaced by the capacitance of the high-voltage busing, using the potential tap, as the means for coupling to the high-voltage circuit, with the effect of the capacitive impedance of the bushing being reduced by an adjustable reactor connected to the bushing tap. The voltage-measuring instrument is described in ANSI C63.2.

Basic-Impulse Insulation Level (BIL) Reduction. The need to demonstrate absence of significant partial discharges in operation is increased for higher circuit voltages where improved surge-arrester characteristics have encouraged a continuing trend toward BIL reduction. Because of progressively decreasing margins between the conventional induced test voltage and operating voltage, new standards for transformers rated 115 kV and above require a 1-h induced voltage test with continuous monitoring of partial-discharge levels to demonstrate the soundness of the insulation. During this test all parts of the insulation system must be overstressed to a degree corresponding to 150% of maximum system voltage at the high-voltage terminals (see ANSI/IEEE C57.12.00).

Partial Discharges in Transformers. This may also be detected by acoustic transducers in the oil or on the tank wall. If a sensitive transducer shows no partial discharges, any partial discharges picked up on the bushing tap originate outside the transformer. If the transducer shows corona, it can be used to locate the source of partial discharges within the transformer tank by measuring the time interval after the partial discharges voltage appears at the bushing tap until the effect appears at the transducer. Then the distance from the transducer to the source of partial discharges is 1 in for each 15 μs of delay.

10.1.10 Radio-Influence Voltage

Excessive partial discharges may cause high-frequency voltages to appear at the terminals which can interfere with radio communication. Suitable maximum limits of voltage in compliance with Federal Communication Commission requirements have been established and are shown in NEMA Publication TR 1. For power transformers, this limits the high-frequency voltage at 1 MHz, measured at about 110% of operating voltage, to 250 μV up to 14.4 kV operating, 650 μV up to 34.5 kV operating, 1250 μV up to 69 kV operating, and 5000 μV up to 345 kV operating.

10.1.11 Testing

Standard Tests. ANSI/IEEE C57.12.00 defines routine, design, and other tests for liquid-immersed transformers. The following are listed as routine tests for transformers 501 kVA and larger:

1. Measurement of resistances of the windings
2. Measurement of turns ratio

3. Phase-relation tests: polarity, angular displacement, and phase sequence
4. No-load loss and exciting current
5. Load loss and impedance voltage
6. Low-frequency dielectric tests (applied voltage and induced voltage)
7. Leak test on the transformer tank
8. Lightning-impulse tests (full wave and chopped wave; for transformers with high-voltage windings from 115 through 765 kV)

The following are listed as design tests for transformers 501 kVA and larger (required on only one unit of a given design):

1. Temperature rise tests (could be omitted if a unit which is essentially a thermal duplicate had been previously tested)
2. Lightning-impulse tests (full wave and chopped wave; for transformers with high-voltage windings of 69 kV and below)
3. Audible sound level
4. Mechanical tests of lifting and moving devices
5. Pressure test on the transformer tank

Other tests listed in ANSI/IEEE C57.12.00 (including short-circuit tests and specialized dielectric tests) shall be made only when specified. Test procedures for all routine and design tests (and many other tests) are defined in the test code document ANSI/IEEE C57.12.90.

The regulation of a transformer may be determined by loading it according to the required conditions at rated secondary voltage and measuring the rise in secondary voltage when the load is disconnected. The rise in voltage when expressed as a percentage of the rated voltage is the percentage regulation of the transformer. This test is seldom made, because the regulation is easily calculated from the measured impedance characteristics.

Efficiency of a transformer is seldom measured directly, because the procedure is inconvenient and the efficiency can be readily calculated.

10.1.12 Oil-Preservation Systems and Detection of Faults

Oil-Preservation Systems. Although transformer oil is a highly refined product, it is not chemically pure. It is a mixture principally of hydrocarbons with other natural compounds which are not detrimental. There is some evidence that a few of these compounds are beneficial in retarding oxidation of the oil.

Although oil is not a “pure” substance, a few particular impurities are most destructive to its dielectric strength and properties. The most troublesome factors are water, oxygen, and the many combinations of compounds which are formed by the combined action of these at elevated temperatures. A great deal of study has been given to the formation of these compounds and their effects on the dielectric properties of oil, but there apparently is no clear relation between these compounds and the actual dielectric strength of the transformer insulation structure.

Oil will dissolve in true solution a very small quantity of water, about 70 ppm at 25°C and 360 ppm at 70°C. This water in true solution has relatively little effect on the dielectric strength of oil. If, however, acids are present in similar amounts, the capacity of oil to dissolve water is increased, and its dielectric strength is reduced by the dissolved water. Small amounts of water in suspension cause severe decreases in dielectric strength. The primary reason for concern over moisture in transformer oil, however, may not be for the oil itself but for the paper and pressboard which will quickly absorb it, increasing the dielectric loss and decreasing the dielectric strength as well as accelerating the aging of the paper.

It is generally recognized today that the best answer to the problem of air and water is to eliminate them and keep them out.

For this purpose, in American practice, transformer tanks are completely sealed. About three basic schemes are used in sealed transformers to permit normal expansion and contraction of oil (0.00075 per unit volume expansion per degree Celsius) as follows:

1. A gas space above the oil large enough to absorb the expansion and contraction without excessive variation in pressure. Some air may unavoidably be present in the gas space at the time of installation but soon the oxygen mostly combines with the oil without causing significant deterioration, leaving an atmosphere which is mostly nitrogen.
2. A nitrogen atmosphere above the oil maintained in a range of moderate positive pressure by a storage tank of compressed nitrogen and automatic valving. This scheme has the advantage that the entrance of air or moisture is prevented by the continuous positive internal pressure, and the disadvantage of somewhat higher cost.
3. A constant-pressure oil-preservation system consisting of an expansion tank with a flexible synthetic-rubber diaphragm floating on top of the oil. This scheme has the advantages that the oil is never in contact with the air and there is always atmospheric pressure and not a variable pressure on the oil. The disadvantage is the higher cost. A number of mechanical variations and elaborations of this general idea have been devised.

It is now generally recommended that the constant-pressure oil-preservation system of item 3 be employed on all high-voltage power transformers (345 kV and above) and on all large generator step-up transformers. This is a consequence of unfavorable experience with transformers having gas-cushion systems, which inherently operate with large quantities of the cushion gas in solution in the hot oil under load. If the oil is suddenly cooled (reduction of ambient temperature or load), the oil volume contracts and the static pressure of gas over the oil drops rapidly, allowing free gas bubbles to come out of solution throughout the insulation system. The dielectric strength of the oil and cellulose insulation system is drastically weakened when it has free gas inclusions, and this has occasionally led to electrical failure of operating transformers.

Fault Detection. Detection of internal faults in transformers at an early stage of their development is most desirable to limit the extent of damage. Two levels of seriousness of faults are recognized. Incipient (or developing) faults have not yet progressed to the point where they affect the functional capability of the transformer, but it is likely that their seriousness will increase with time if not corrected. Examples would include partial discharge sites within the insulation, intermittent low-energy sparking, overheated conductor insulation, or hot metal parts in contact with oil only. More serious or permanent internal faults affect functionality immediately and must be removed quickly before their consequences can jeopardize the safety of personnel or other equipment.

Most commonly employed means of sensing incipient faults relate to detection of gases generated at the fault site. Automatically operating gas-detection devices which can be supplied on the transformer employ any of the following principles:

1. Free gas accumulation at the cover
2. Sensing of combustible gases within a gas cushion over the oil
3. Separation of certain gases dissolved in the oil

These devices are indicators of possible internal problems. Verification of the problem should be done by gas-in-oil analysis using a gas chromatograph. In addition, periodic manual sampling of the oil for laboratory analysis should be practiced. The composition of the gas dissolved in the oil is very useful for diagnosis of the nature of an incipient fault.

Permanent internal faults can be detected by fault-pressure relays or differential relays, either of which give a signal that can be used to trip circuit breakers and remove the transformer from the system. The fault-pressure relay senses the sudden buildup of pressure produced by arc-generated gases after a fault has occurred. Unfortunately, such relays can also be operated by any other event which causes a rapid pressure change, so they cannot be set to be too sensitive. Differential relays sense that more current is flowing into the transformer than is flowing out, but relays which are insensitive to the initial inrush of exciting current should be used. After a transformer has been disconnected as a result of relay operation, it is always desirable to get it back into service as quickly as possible. Following differential-relay operation, circuit breakers may be reclosed to check whether the fault is self-healing. The penalty for reconnection of a damaged transformer is that if the fault recurs, the damage to the transformer and possibly associated equipment will be greater. Under no circumstances should a transformer be reconnected to the system following operation of a fault-pressure relay without thorough investigation of the cause of the relay operation.

When a transformer has been taken out of service because of fault indications, the following procedure should be used:

1. If there is a gas space, take samples of the gas from the gas space for analysis to determine whether products of decomposition are present.
2. Take oil samples for extraction of dissolved gases for similar analysis.
3. Make insulation power factor, insulation resistance, and turns ratio tests to check whether their results conform to normal values.
4. Perform any other tests which seem to be indicated by the results of the first tests.
5. Check the operation and calibration of the protective relay.

10.1.13 Overcurrent Protection

Effects of Overcurrent. A transformer may be subjected to overcurrents ranging from just in excess of nameplate rating to as much as 10 or 20 times rating. Currents up to about twice rating normally result from overload conditions on the system, while higher currents are a consequence of system faults. When such overcurrents are of extended duration, they may produce either mechanical or thermal damage in a transformer, or possibly both. At current levels near the maximum design capability (worst-case through-fault), mechanical effects from electromagnetically generated forces are of primary concern. The pulsating forces tend to loosen the coils, conductors may be deformed or displaced, and insulation may be damaged. Lower levels of current principally produce thermal heating, with consequences as described later on loading practices. For all current levels, the extent of the damage is increased with time duration.

Protective Devices. Whatever the cause, magnitude, or duration of the overcurrent, it is desirable that some component of the system recognize the abnormal condition and initiate action to protect the transformer. Fuses and protective relays are two forms of protective devices in common use. A fuse consists of a fusible conducting link which will be destroyed after it is subjected to an overcurrent for some period of time, thus opening the circuit. Typically, fuses are employed to protect distribution transformers and small power transformers up to 5000 to 10,000 kVA. Traditional relays are electromagnetic devices which operate on a reduced current derived from a current transformer in the main transformer line to close or open control contacts, which can initiate the operation of a circuit breaker in the transformer line circuit. Relays are used to protect all medium and large power transformers.

Coordination. All protective devices, such as fuses and relays, have a defined operating characteristic in the current-time domain. This characteristic should be properly coordinated with the current-carrying capability of the transformer to avoid damage from prolonged overloads or through faults. Transformer capability is defined in general terms in a guide document, ANSI/IEEE C57.109, *Transformer Through Fault Current Duration Guide.* The format of the transformer capability curves is shown in Fig. 10-35. The solid curve, *A*, defines the thermal capability for all ratings, while the dashed curves, *B* (appropriate to the specific transformer impedance), define mechanical capability. For proper coordination on any power transformer, the protective-device characteristic should fall below both the mechanical and thermal portions of the transformer capability curve. (See ANSI/ IEEE C57.10-38 for details of application.)

10.1.14 Protection Against Lightning

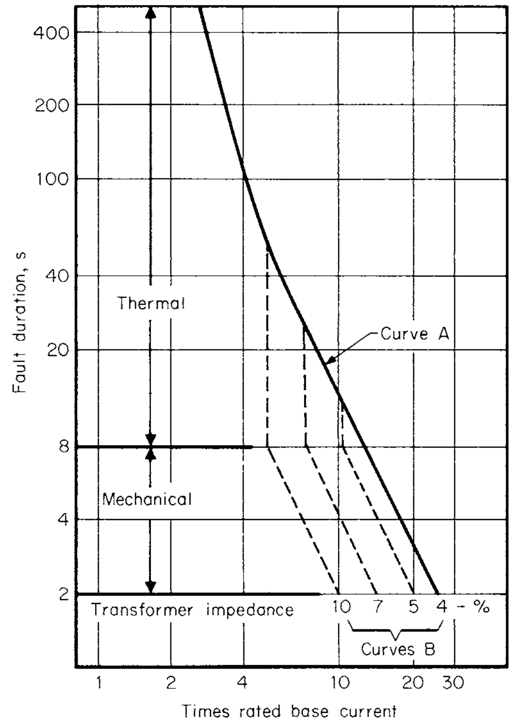


FIGURE 10-35 Transformer through-fault protection curves.

Impulse insulation level may be demonstrated by factory impulse-voltage tests using 1.5 × 50-μs full waves and chopped waves. The full wave demonstrates the BIL for traveling waves coming into the station over the transmission line. The chopped wave demonstrates strength against a wave traveling along the transmission line after flashing over an insulator some distance away from the transformer. These waves do not simulate direct lightning strokes on or near the transformer terminals, which would result in the application of a steep-front wave to the transformer winding. Such strokes are usually avoided by ground wires or protecting grounded structures.

A transformer may be subjected to severe lightning voltages as a result of a direct stroke to the transformer terminal, adjacent bus, or transmission line. Less severe voltages may result from strokes on a distant part of the system or from strokes to ground near the system. Since lightning voltage may exceed the insulation strength of the transformer, protection is necessary.

Voltage-time curves are used in evaluating protection, because for short times the insulation strength changes significantly with duration of voltage. Protection is effective if the voltage-time curve of the transformer is above the voltage-time curve of the protective equipment, so that for any time duration the kilovolts insulation strength of the transformer exceeds the protective level at the same duration. The voltage-time curves of transformer insulation have considerable "turn-up," that is, for durations under 10 μs the kilovolts insulation strength is much greater. Rod gaps in air are unsuitable for protecting transformers, because they have even more turn-up than do transformers.

Surge Arresters. The modern surge arrester has very little turn-up and is an essential adjunct to the transformer whenever there is lightning exposure.

The required surge arrester rating depends on the effectiveness of the neutral grounding. The rating is expressed in percent of rated line-to-line power-frequency voltage that the arrester will withstand. Effectiveness of system grounding is described by the ratios of the zero-sequence resistance and impedance to the positive-sequence resistance and impedance. An 80% arrester is commonly used when the ratio of zero sequence to positive sequence is between 0.5 and 1.5 for resistance and between 1 and 3 for impedance. Lower ratios may permit 75% or 70% arresters. Higher ratios may require 85% or 90% arresters. Use of the 100% protective level is not economical at high voltages (Honey 1987).

Figure 10-36 shows the voltage-time curve of a 345-kV transformer, with 1050-kV BIL (reduced by two steps) compared with the voltage-time curve of a 90% metal-oxide surge arrester. Since this new type of arrester has no series gaps to spark over, the characteristic is described as a protective

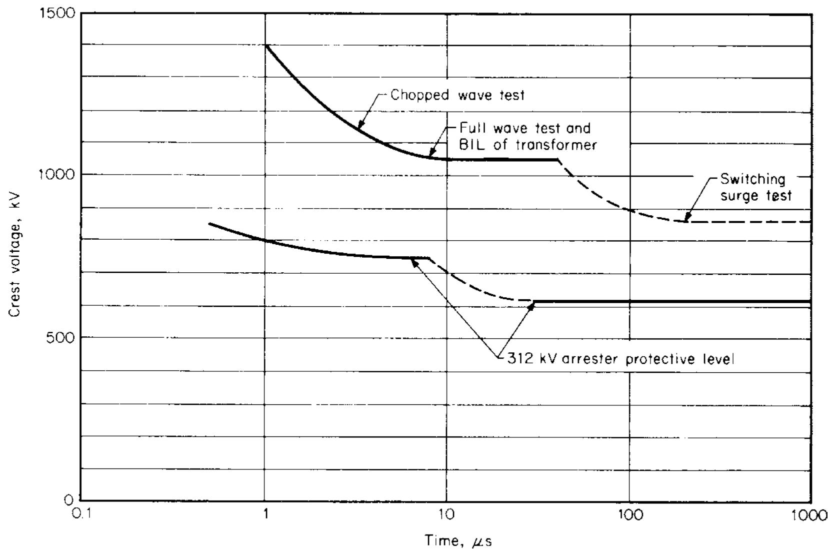


FIGURE 10-36 Surge-arrester protection margin for impulse and switching-surge conditions for a 345-kV transformer with BIL reduced two steps. The arrester curve shown is for an 8 × 20 μs current wave of 15-kA crest.

level rather than a sparkover level. The voltage-time curve of an arrester depends on the amount of impulse current. The voltage-time curve shown in Fig. 10-36 corresponds to 15,000-A crest, which is not likely to be exceeded except by direct strokes. For best protection, the arrester should be located as close as possible to the terminals of the transformer, and the arrester ground should be connected by a short, direct conductor to the transformer tank and substation grounding system.

The steep rate of rise of lightning current as such may result from a nearby direct stroke will produce discharge voltages higher than shown in Fig. 10-36 and may damage the lightning arrester. Therefore, the substation and the first half mile of connected transmission lines should be protected against direct strokes by a suitable combination of grounded masts and ground wires (Linck 1975).

10.1.15 Installation and Maintenance

Proper installation and maintenance of power transformers is needed to assure long life in service. General requirements from ANSI C57.93 and from manufacturer's instructions are summarized below. Most power transformers have additional manufacturer's installation and maintenance instructions which should be carefully observed. These instructions, including safety practices, should be followed for the protection of workers and the transformer.

Transformers are processed tested, and inspected before shipment. Smaller transformers are shipped filled with oil and fully assembled when shipping clearances and weights permit. For most large-power transformers, the external components, such as coolers and bushings, must be removed to meet shipping dimensions. Also, the oil is removed to reduce weight, and the unit is secured in dry gas.

Inspection on Arrival. Before removal from the car, inspect for shipping damage. If damage is found, a claim should be filed with the carrier and the manufacturer should be notified.

Transformers shipped oil-filled should be inspected for evidence or leakage or entrance of moisture during shipment. If the transformer is received in damaged condition, tests should be made to check the transformer for dryness. Oil samples taken from the bottom valve should be tested for moisture limits. Where bushings are shipped in place, insulation power-factor measurements can be used. Power factors of new transformers range from 0.2% to 0.5% at 25°C. In any case, a power-factor measurement may be helpful for comparison with test values recorded by the manufacturer prior to shipment.

Transformers shipped gas-filled are fitted with a connection to which a compound pressure/vacuum gage can be connected. The gage should have a range of +10 lb/in^2. A positive or negative pressure indicates that the tank is tight; a continuous zero reading indicates a probable leak. If there is a reason to question dryness, the moisture can be estimated by a dew-point measurement of the gas in the transformer. Add dry gas immediately to about 3-lb/in^2 (gage) pressure and check for leaks.

Dew point is sensitive to temperature changes; therefore it is essential that the insulation temperature be accurately determined. Some instruments will not provide accurate dew-point readings below 32°F. The final degree of insulation dryness should be determined by measuring the insulation power factor or insulation resistance. If the measured dew point exceeds an acceptable value, contact the manufacturer for recommended course of action.

A preliminary internal inspection is not normally required unless shipping damage is apparent. (*Caution:* Do not enter any transformer until the gas in the tank is replaced by dry air with at least 19.5% oxygen content. Exercise caution when entering a transformer and avoid introducing contamination. The time the transformer is open for inspection should not exceed 2 h, and dry air should be circulated.)

Oil sampling. Samples of oil should be taken from the bottom. An oil-sampling valve is provided at the bottom of the transformer tank for this purpose. A metal or glass thief tube can be conveniently used to obtain a bottom sample from an oil barrel. Test samples should be taken only after the oil has settled for some time, varying from 8 h for a barrel to several days for a large transformer. Cold oil is much slower in settling. In drawing samples of oil from a sampling valve, some oil should first be discarded so that the sample will come from the bottom of the container and not from the sampling pipe. Examine a sample in a clear glass container for free water, which in any quantity is readily observable. The sample container should be a large-mouthed glass bottle, 1 qt or larger, with cork or glass stopper. The bottle should be carefully cleaned and dried before being used. Bottles should be of amber color if samples are to be stored to be tested later for color or sludge-forming characteristics. Refer to ASTM 923-81 for important details of oil-sampling technique.

Testing for Oil Dielectric Strength. The testing fixture should be cleaned thoroughly to remove any particles or fibers and rinsed out with a portion of the oil to be tested. The testing fixture should be filled with oil, and both oil and fixture should be at room temperature. Allow 3 min for air bubbles to escape before applying voltage. Tests are made by two methods. ASTM D877 uses 1-in-diameter square-edge electrodes spaced 0.10 in apart and a rate of voltage rise of 3000 V/s. ASTM D 1816 uses special radiused-surface electrodes spaced 0.04 in apart, with continuous oil circulation, and a rate of voltage rise of 500 V/s. The latter test is more sensitive to slight moisture or particulate contamination. In either case, the average voltage for five breakdowns is taken as the dielectric strength of the oil. Strength of new oil should exceed the minimum value for good oil as shown in Table 10-4. (See also ANSI/IEEE C57.106.)

TABLE 10-4 Dielectric Strength of Oil

kV average dielectric strength by ASTM D877-82	kV average dielectric strength by ASTM D1816-82	Condition of oil
30 or over	29 or over	Good
26 to 29	23 to 28	Usable
Under 26	Under 23	Poor

Filtering to Increase Dielectric Strength. If the oil tests below "good," it should be filtered to remove impurities and moisture. It is best to discharge filtered oil into a clean, dry tank and avoid mixing with unfiltered oil. If the filtered oil must be discharged back into the transformer tank, the oil should be withdrawn from the bottom filter-press valve and, after filtering, returned through the top filter-press valve. Oil should not be filtered while the transformer is energized, because the dielectric strength may be temporarily reduced by aeration. If no facilities are available for making dielectric tests, send a sample to the manufacturer marked with the serial number of the transformer.

Drying the Core and Coils. This should be necessary only if an accident occurs during shipping, storage, or service. (Refer also to ANSI/IEEE C57.12.11 and C57.12.12.)

If possible, determine the extent of the moisture and manner in which it entered the tank. The manufacturer should be contacted for recommendations concerning additional checks and strips for drying out the transformer. If drying is necessary, one or more of the following methods may be used, depending on the facilities available:

1. Heat and vacuum
2. Vacuum only
3. Heat in oil
4. Heat in air

A low value of residual moisture can be attained most rapidly by method 1. Method 2 is effective but requires longer time and better vacuum equipment. Method 3 is very slow and not as effective as the vacuum method. Method 4 is recommended only for smaller, low-voltage transformers.

Method 1: Heat and Vacuum. The following procedures may be used to elevate the temperature of the core and windings prior to the vacuum process:

1. Using external heaters and pump equipment, spray hot oil through the cover of the transformer. Maintain a vacuum of 10 torr or less on the transformer during the oil spraying operation in order to prevent oxidation of the oil and to aid in the removal of gas from the insulation. Pump the oil from the bottom of the transformer through filters, through the heat exchanger, and back to the cover of the transformer. Use the minimum amount of oil necessary to establish circulation. Degassing-dehumidifying equipment may be used if the equipment is available. The oil temperature entering the top of the transformer should be between 50°C (122°F) and 75°C (167°F). The temperature of the core and coils will be elevated to equilibrium conditions when the output oil temperature becomes constant, and the temperature of the core and coils will be near the temperature of the output oil.
2. Heat may also be generated by circulating current through the windings. This requires a controlled source of power connected to either winding, with the other winding short-circuited. For this method, the transformer should be filled to normal level. Connect the vacuum pump to a suitable valve above the oil level. The temperature of the windings should not exceed 95°C and the oil 85°C. Initial currents up to full-rated current may be used, then reduced as the temperature approaches the desired value. The voltage required to circulate this current will depend on the impedance of the transformer, which can be obtained from the instruction plate. After the desired winding and oil temperatures have been reached, disconnect the power supply and drain the oil from the transformer tank. When the oil has been drained from the tank, close the valves and start the vacuum pump. Continue pulling vacuum on the transformer until the water extraction stops.

Method 2: High Vacuum. This method requires the use of a suitable vacuum pump, capable of pumping down to an absolute pressure of 0.05 mmHg (6.67 Pa) or lower, and a refrigerated vapor trap to collect the water. No additional heat is required if an adequate vacuum pump is used.

Drain the oil from the transformer, filling the tank with dry nitrogen as the oil is drained. Remove heat exchangers and other external pipe connections and seal these openings, preferably with blanking plates. Connect the vapor trap and vacuum pump to a suitable pipe connection on the tank. Seal the tank and pressure test for leaks. After ensuring that all leaks have been eliminated, start the vacuum pump. Water extraction from the insulation will begin when the residual vapor pressure in the tank is reduced below the vapor pressure of water in the insulation. Drying may be continued as long as moisture is being extracted or may be terminated when the residual moisture content of the insulation has been reduced to the desired level as determined from a moisture equilibrium chart.

Method 3: Heat in Oil. This method is slow and limited to smaller, low-voltage transformers. Also, method 3 may be used for transformer tanks not designed for full vacuum. This method requires the use of a suitable oil filter with either a vacuum-drier-type or blotter press, plus a heater which will enable hot oil at a temperature of approximately 85°C to be circulated in the transformer tank until positive indication of drying out of the windings has been obtained. The transformer should be filled with oil to cover the core and windings and the oil circulated through either the filter-press valve and drain valve or through top and bottom radiator valves. Wherever possible, to reduce heat losses due to radiation, prevent the oil from circulating through the coolers. Blanket the outside of the transformer tank to reduce to a minimum the time of drying out and the amount of heating required to keep the oil temperature constant.

With this method the moisture is removed through the oil filter. If a blotter-press filter is used, the rate of water extraction will depend on the degree of saturation of the filter papers. Filter papers must be extremely dry and papers must be changed frequently if this method is to be effective. The rate of transfer of moisture increases with temperature; hence it is desirable to operate at the highest temperature which will not cause deterioration of the oil.

Method 4: Hot-Air Drying. This is limited to small low-voltage transformers and where clean dry air is available. With the transformer assembled in its tank, the tank should be blanketed in order to reduce to a minimum the amount of heating required, and also to keep the interior of the tank at a uniform temperature to prevent condensation in the tank.

Dry air should be forced by a fan over the heating elements, then through an opening at the base of the tank to pass over and through the coils before exhausting through an opening in the cover. Baffles should be placed between the hot-air inlet and the windings to prevent the flow of hot air from being concentrated on one small portion of the windings.

Time Required for Drying. This can range from 72 h to 3 weeks depending on the size and condition of the transformer and the method of drying. In general, the use of a high vacuum and a cold trap is faster and more efficient than heat alone.

Insulation Resistance. This will indicate the degree of dryness only when the transformer is dried without oil. If the initial insulation resistance is measured at room temperature, it may be high, although the insulation is not dry, but as the transformer is heated up, it will drop rapidly.

As the drying proceeds at a constant temperature, the insulation resistance will generally increase gradually until toward the end of the drying period, when it increases quite rapidly and then levels off at a high value. The drying should continue until the resistance is constant for a period of 12 h.

Insulation Power-Factor Reading. These readings (at 60 Hz) will indicate the degree of dryness. The power factor will first increase as the temperature increases and then will gradually decrease as drying progresses. Drying should continue until the power factor is constant for a period of 12 h. If power factor is measured on transformers dried in oil by the short-circuit method, the power factor should be used to supplement oil tests as a measure of dryness.

Filling without Vacuum. (*Note:* This method should be practiced only on low-voltage transformers. Check manufacturer's recommendations.) Use extreme care to keep moisture out of the core and coils. The tank should not be opened to the atmosphere until the core and coils are under oil, unless vacuum filling is available. The oil shipping tank or oil drums should not be opened until their temperature is the same as or higher than that of the surrounding air and the transformer is in place and ready to receive the oil. Metal or synthetic rubber hose should be used for filling, because transformer oil is contaminated by natural rubber. Oil should never be added to a transformer without passing through a filter press.

Static charges can be developed when transformer oil flows in pipes, hoses, or tanks. Oil leaving a filter press may be charged to over 50,000 V. To accelerate dissipation of the charge in the oil, ground the filter press, the tank, and all bushings or winding leads during oil flow into any tank. Conduction through oil is slow; therefore it is desirable to maintain these grounds for at least 1 h after the oil flow has ended.

Avoid explosive gas mixtures in any container into which oil is flowing. Arcs can occur along the surface of the charged oil even though all metal is grounded.

Filling with Vacuum. The vacuum line should be connected to a tapped opening on a cover-mounted shipping plate or to a valve near the top of the tank. An opening of 2 in minimum is recommended. The oil line can be connected to a suitable opening on a cover-mounted shipping plate or the top filter-press valve. The oil line should always be connected at the top of the tank so that the oil can be deaerated as it enters.

Transformers with operating voltage less than 161 kV and with core and coils not exposed to the atmosphere should be filled under vacuum better than 25 mmHg absolute pressure. The vacuum should be held 4 h before filling and continued during filling until the core and coils are covered. The vacuum can then be removed for installation of bushings and the remaining oil added without vacuum. Transformers with an operating voltage of 161 kV and above or transformers with core and coils exposed to the atmosphere should be completely filled under a vacuum better than 2 mmHg absolute pressure. A 2-mm vacuum should be held until the tank is filled to the 25°C level.

The filling rate should be under 1500 gal/h to facilitate evacuation and complete oil filling of all air pockets and voids. It is also recommended that the core and coil temperature be above 10°C to prevent “frost” formation.

Energization. When the voltage is first applied, it should, if possible, be brought up slowly to its full value so that any wrong connection or other trouble will be discovered before damage results. After full voltage has been applied successfully, the transformer should preferably be operated for a short period without load.

When the transformer is first energized it should be kept under close observation for the first 8 h. Check and record the oil temperature, the winding temperature, the tank pressure, and the ambient temperature. Watch particularly for any sudden changes. After 7 days of operation, check for oil leaks and for abnormally high usage of nitrogen if the transformer is equipped with Inertaire. Stop all oil and gas leaks. The observation should continue on a daily schedule for 7 days and then weekly for the first month of operation. An oil sample should be taken during the first month of operation for gas-in oil analysis. This analysis should be repeated annually. Heat exchangers, tap changer, pumps, fans, etc., should be serviced per manufacturer’s instructions.

Internal Inspection of In-service Transformers. This is not necessary unless there is a specific indication of a problem. Oil analysis is a good method to discover potential problems. Sludging of the oil, low dielectric strength, moisture in the oil, or the presence of combustible gases are conditions that may merit an internal inspection of the transformer.

Generation of combustible gas usually indicates internal trouble (not necessarily serious). Analysis of the gas sometimes helps to identify the source. If collection of combustible gas continues without discoverable cause, partial discharge voltage measurement may establish whether or not there is an internal fault.

Severe system disturbances, incidence of through-fault, or a circuit-breaker operation would also be reason for an internal inspection of a transformer.

Operating without Cooling. A liquid-cooled transformer should not be run continuously, even at no load, without the cooling liquid. In an emergency, forced-oil air-cooled transformers may be operated without fans and pumps (1) at rated load for approximately 1 h, starting at full-load temperature rise,

(2) at rated load for approximately 2 h, starting cold (at ambient temperature), (3) at rated voltage and no load for approximately 6 h, starting at full-load temperature rise, and (4) at rated voltage and no load for approximately 12 h, starting cold.

When only a portion of the cooling equipment is operating, the transformer may be operated at reduced load approximately as indicated in Table 10-5.

TABLE 10-5 Operation with Limited Cooling Equipment

Percent of cooling equipment in operation	Percent of rated load that may be carried
33	50
40	60
50	70
80	90

10.1.16 Loading Practice

The *temperature limitation of loading* must be considered. Ordinarily, the kVA that a transformer should carry is limited by the effect of reactance on regulation or by the effect of load loss on system economy. At times, it is desirable to ignore these factors and increase the kVA load until the effect of temperature on insulation life is the limiting factor. High temperature decreases the mechanical strength and increases the brittleness of fibrous insulation, making transformer failure increasingly likely, even though the dielectric strength of the insulation material may not be seriously decreased. Overloading of transformers should be limited by reasonable consideration of the effect on insulation life and the probable effect on transformer life.

The *insulation life of a transformer* is defined as the time required for the mechanical strength of the insulation material to lose a specified fraction of its initial value. Loss of 50% of the tensile strength is the usual basis for evaluating conductor insulation for power transformers. (*Note:* A transformer may continue to perform beyond the predicted life if it is not disturbed by short-circuit forces, etc. This is because at the mechanical end point, the disintegration of the fibers is not total as shown by degree of polymerization measurements, the true indicator of paper insulation integrity.)

The *aging of insulation* is a chemical process that occurs more rapidly at higher temperatures according to the Arrhenius reaction-rate theory, as expressed in

$$\log h = \frac{K_1}{C + 273} + K_2 \tag{10-63}$$

where C = temperature in degrees Celsius of insulation, K_1, K_2 = constants determined by test, and h = hours of life. Use of this equation permits results of relatively short duration tests at relatively high temperature to be extrapolated to indicate probable insulation life at moderate temperatures. ANSI/IEEE C57.92-1981 contains loading recommendations for power transformers up to 100 MVA with 55°C and 65°C average winding-rise insulation systems based on extrapolated life tests. Figure 10-37 shows the corresponding curves of rate of loss of life as a function of temperature as defined in this document. The constants in Eq. (10-63) are

$$55°\text{C rise} \qquad K_1 = 6972.15 \qquad K_2 = -14.133$$

$$65°\text{C rise} \qquad K_1 = 6972.15 \qquad K_2 = -13.391$$

ANSI/IEEE C57.91 provides similar loading recommendations for distribution transformers, including the following values for the constants in Eq. (10-63):

$$55°\text{C rise} \qquad K_1 = 6328.8 \qquad K_2 = -11.968$$

$$65°\text{C rise} \qquad K_1 = 6328.8 \qquad K_2 = -11.269$$

Accepted methods for functional life evaluation have been established for both distribution transformers and for power transformers.

To determine the aging of the insulation resulting from a specific daily load cycle (1) establish an approximately equivalent stepped load cycle, (2) calculate the resulting curve of hot-spot

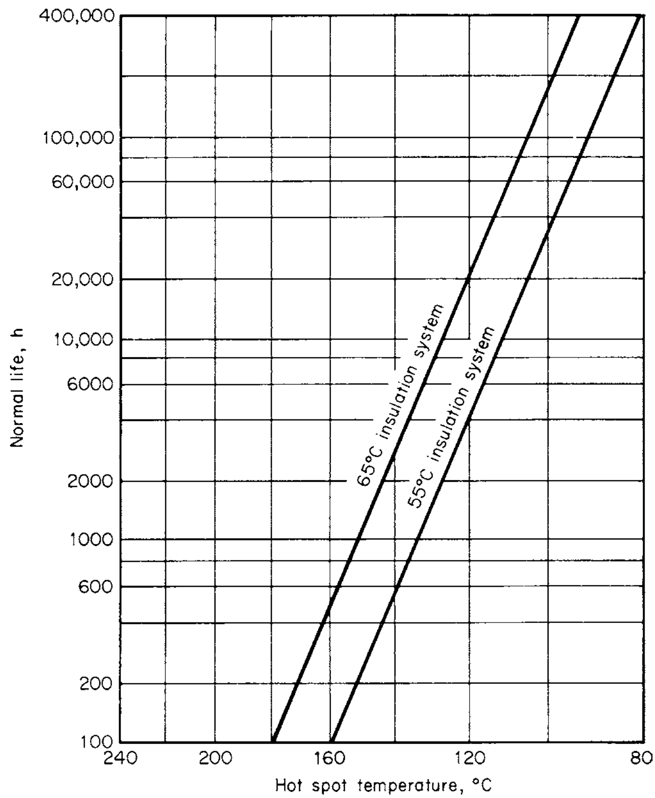


FIGURE 10-37 Loss of life as a function of temperature for power transformers up to 100 MVA under loading conditions of ANSI/IEEE C57.92.

temperature, (3) replace the hot-spot temperature curve by an approximately equivalent stepped curve, (4) calculate the percent aging for each step from the applicable curve of Fig. 10-37, and (5) add the aging for all the steps in the daily cycle. The result is the fraction of insulation life used up each day. The reciprocal is the number of days of total insulation life if the same load cycle repeats every day.

For example, consider a transformer with a daily load cycle of 4 h at 140% load and 20 h at 80% load in 30°C ambient. The hot-spot temperature curve shown in Fig. 10-23 is reproduced in Fig. 10-38, together with an equivalent stepped curve. The calculation of loss of life per day is shown in Table 10-6.

The normal life at every temperature can be determined from Fig. 10-37 or from Eq. (10-63) for the 65°C rise insulation system. The fraction of the life consumed during each time step can then be calculated and summed for all of the time steps in the 24-h period. In this case, 0.09% of the life is consumed in one day, so the total life would be 1111 days or about 3 years if this load cycle were continued. For comparison, a transformer with a 65°C average winding-rise insulation system (80°C hot-spot rise) operating in a 30°C ambient would have a hot-spot temperature of 110°C and a normal life of 65,000 hours or 7.4 years. The shortening of the insulation life from 7.4 years to 3 years is a measure of the severity of the load cycle. The actual transformer life may, of course, be shorter or longer, depending on exposure to overvoltage, overcurrent, shock, contamination, fault conditions, etc.

Loading-capability tables for normal consumption of life and for moderate sacrifice of life are documented in ANSI/IEEE C57.92 for power transformers. Information is provided for situations

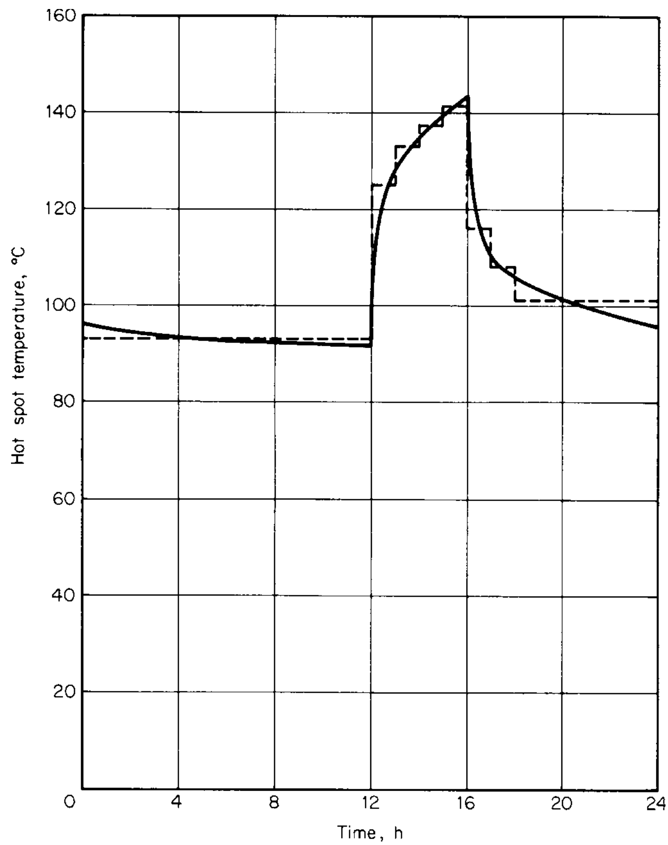


FIGURE 10-38 Equivalent stepped curve of hot-spot temperature for loss-of-life calculations of a daily load curve.

involving different ambient temperatures, different peak-load durations, and different loads prior to the peak for an assumed set of representative transformer characteristics.

Ambient temperature affects load capacity by an amount dependent on the type of cooling, as shown in Table 10-7.

TABLE 10-6 Calculation of Loss of Life per Day on Daily Load Cycle

Duration of step, h	Temperature of hot spot, °C	Life, h, at temperature	% loss of life on step
12	93	455,600	0.003
1	125	13,396	0.007
1	133	6,050	0.017
1	137	4,114	0.024
1	141	2,818	0.035
1	116	34,062	0.003
1	108	81,023	0.001
6	101	178,284	0.003
			0.093

TABLE 10-7 Effect of Ambient Temperature on kVA Capacity

Type of cooling	% of rated kVA decrease in capacity for each °C increase over 30°C air or 25°C water	% of rated kVA increase in capacity for each °C decrease under 30°C air or 25°C water
Self-cooled*	1.5	1.0
Water-cooled†	1.5	1.0
Forced-air-cooled*	1.0	0.75
Forced-oil-cooled*	1.0	0.75

*From 0 to 50°C air temperature.
†Up to 35°C water temperature.

For ambient temperature of air-cooled transformers, use the average value over a 24-h period or 10°C under the maximum temperature during the 24-h period, whichever is higher. For ingoing water temperature, use the average value over a 24-h period or 5°C under the maximum temperature during the 24-h period, whichever is higher.

Limitations. The temperature of the top oil should never exceed 110°C for power transformers with a 55°C average winding-rise insulation system or 110°C for those with a 65°C average winding-rise insulation system. The consequence of exceeding these limits could be oil overflow or excessive pressure. The winding hot spot should not exceed 150°C for the 55°C average winding-rise insulation system or 180°C for the 65°C average winding-rise insulation system. These limitations are based principally on a concern for rates of insulation aging, but it should be noted that free bubbles may be evolved at hot-spot temperatures above 140°C, with consequent weakening of dielectric strength. The peak short-duration loading should never exceed 200% of rating, except for transformers rated over 100 MVA a limit of 150% of rating is recommended (IEEE Standard 756). This reflects a concern for stray flux heating in large units.

10.1.17 Loss Evaluation

Loss evaluation is a procedure by which the buyer and seller achieve an economic balance in adding material to the transformer design to get lower losses. It is achieved by establishing a value in dollars per kilowatt for load loss and a similar value for no-load loss.

An incremental investment in capacity is required to generate power to supply loss and bring it to the transformer. In addition, there is a continuing expense for fuel to supply the lost power. The continuing expense is converted to present worth and added to the incremental investment to give the total present worth of the loss. This present worth of a kilowatt of loss is naturally higher for the no-load loss, which is continuous, than it is for the load loss, and the value is higher the farther the transformer is from the generator; the values will of course depend on the accounting rules and procedures in force at the particular location.

The following equations are commonly used to establish loss evaluations:

$$V_L = S + \frac{8760EF_L}{R} \tag{10-64}$$

$$V_N = S + \frac{8760EF_N}{R} \tag{10-65}$$

where E = dollars per kilowatthour cost of energy (this can conceivably be very low for a hydro station but can range up to 0.02 or more for fuel-fired stations, depending on fuel cost, and, of course, the figure will be even higher at locations remote from the generating station), F_L = ratio of average load loss to rated load loss, F_N = ratio of average no-load loss to rated no-load loss (1.00 for continuous operation), R = per unit (%/100) annual carrying charge on system investment (covers insurance, taxes, depreciation, and return on investment), S = dollars per kilowatt system investment

(200 and up, depending on the system investment out to the transformer location), V_L = dollars per kilowatt evaluation of rated load loss, and V_N = dollars per kilowatt evaluation of rated no-load loss.

Since the load losses of a transformer vary as the square of the load, it is important to state the MVA rating at which the load losses will be evaluated. Since it is common practice of most transformer manufacturers to optimize the design of the transformer at its self-cooled rating, the dollar value of losses for the load loss should be specified at the self-cooled rating. If the dollar value of losses for the load loss is specified at some load other than the self-cooled rating, it can be adjusted to the self-cooled rating by multiplying the dollar value by the square of the ratio of the load at which the losses will be evaluated and the self-cooled rating.

It is also important that the transformer manufacturer knows if the buyer is using the present-worth, the levelized annual cost, or the capitalized cost method of evaluation. If the present worth method is being used, the present-worth multiplier should be stated; if the levelized annual cost method is being used, the carrying charge should be stated so the manufacturer, in either case, knows how the dollar values of losses equate to the first cost of the transformer.

Loss evaluation is an important factor in purchasing new transformers, as in many cases the evaluation of the total loss equals or exceeds the price of the transformer.

10.1.18 Autotransformers

Part of an autotransformer winding is common to both primary and secondary circuits. The common portion is called the common winding, and the remainder is called the series winding. The high-voltage terminal is called the series terminal, and the low-voltage terminal is called the common terminal. Part of the power passes from one winding to the other by transformation, and the rest passes directly through without transformation. Figure 10-39 shows an autotransformer compared with an equivalent two-winding transformer. Both have the same ratio of secondary voltage to primary voltage, T, and both have the power output. The fraction $1 - T$ of the power is transformed, and the fraction T passes through without transformation. The fraction $1 - T$, called the "co-ratio," is a measure of the required size of the core and coils as compared with a two-winding transformer. In addition, the losses and reactance are reduced in approximately the same ratio. For a low value of $1 - T$ the economy of an autotransformer compared with a transformer is attractive.

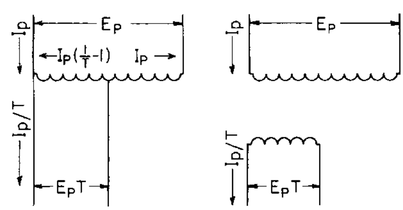


FIGURE 10-39 Comparison of an autotransformer with a two-winding transformer.

The following special characteristics of autotransformers may need consideration: A metallic connection exists between the primary and secondary circuits; this is generally of little consequence with low-voltage circuits, but with high-voltage systems the neutral point must be grounded for safe operation.

The impedance of an autotransformer is normally lower than that of the equivalent two-winding transformer, and the short-circuit current is higher.

Taps near the neutral are relatively ineffective, because turns tapped out at the neutral come out of both circuits and increase core flux density without greatly changing the voltage ratio. The high-voltage rating is more effectively adjusted by taps in the series winding. The low-voltage rating is more effectively adjusted by tapping turns out of the common winding while turns are added to the series winding or by placing the tapping winding in the low-voltage line.

Inversion of the neutral may occur under abnormal conditions in Y-connected autotransformers with ungrounded neutrals. If the voltage on a series terminal is lower than the voltage on the corresponding common terminal, high voltage tends to appear on the neutral. Inversion of the neutral can occur on power-frequency voltage or on transient voltage. Grounding the autotransformer neutral, use of a delta tertiary, and use of three-leg 3-phase cores all help to prevent inversion of the neutral.

Autotransformers are most commonly used to connect two transmission systems at different voltages, frequently with a delta tertiary winding. It is also possible to apply an autotransformer as a generator step-up transformer when it is desired to feed two different transmission systems. In this case the delta tertiary winding is a full-capacity winding connected to the generator, and the two transmission systems are connected to the autotransformer windings. Advantages of the autotransformer when compared to a normal transformer include lower impedance, lower losses, better regulation, smaller size, and lighter weight.

10.1.19 Distribution Transformers

Distribution transformers are generally considered as transformers of 500 kVA, and smaller 67,000 V and below, both single-phase and 3-phase. Older installations are primarily pole-/platform-mounted units. Newer installations are frequently pad-mounted units. Typical applications are for supplying power to farms, residences, public buildings or stores, workshops, and shopping centers.

Distribution transformers have been standardized as to high- and low-voltage ratings, taps, type of bushings, size and type of terminals, mounting arrangements, nameplates, accessories, and a number of mechanical features, so that a good degree of interchangeability results for transformers in a certain kVA range of a given voltage rating. They are now normally designed for 65°C rise.

The most popular primary voltages are 12,470Y/7200, 13,200Y/7620, and 12,000 V delta.

Many of the 2400- and 4800-V primary systems have been converted to 7200 and 7620 V. There is also increasing use of higher-voltage distribution systems such as 24,900Y/14,400 and 34,500Y/19,900 V. Secondary voltage for pole-type units is usually 120/240 or 240/480, 240/120 on single-phase pad-mount units and 208Y/120 on 3-phase units.

Magnetic cores, in general, are composed of cold-rolled silicon steel strip. They take various forms, all designed so that the magnetic flux will pass through the sheet in the direction of the rolling in order to secure the maximum benefit of the superior magnetic quality of this material. For an appreciable portion of the 24-h day, the typical distribution transformer (particularly the pole-mounted 5- to 167-kVA range) is lightly loaded. Because of this, the loss in the core is a significant portion of the total daily loss. Cores for these units are therefore designed for low exciting current and for relatively low core loss to minimize the operating cost. Low-loss cold-rolled silicon strip has contributed materially to reduced losses, weights, and dimensions. New amorphous steel alloys are providing additional opportunities to reduce the units operating costs.

Coils are usually wound in a concentric layer arrangement, with cooling ducts distributed periodically between the layers to maintain reasonable differentials between oil temperature and the average coil and hot-spot temperatures. As a matter of practical operating procedure, distribution transformers are subjected to considerable numbers of overloads for short time periods. Hot-spot temperatures must be limited on these overload excursions if the transformer is to have a long insulation life. It is now general practice to employ thermally upgraded materials in the insulation system to improve aging characteristics. Increased mechanical strength is achieved by using a special intermittent coating of a heat-reactive adhesive on the layer insulation to bond the coil into a rigid mass during the drying process. Heat and vacuum drying plus vacuum oil filling imparts good dielectric strength to the windings.

Aluminum and copper conductors are both used in the coils of distribution transformers. The decision to use one material over another is based on the required loss performance levels for individual utilities. Aluminum conductor is widely used in the secondary windings, where full-width aluminum strip is employed. Such coils are also mechanically stronger.

To cool the unit, the radiating surface of the tank itself suffices in the smaller ratings. In the larger ratings, auxiliary cooling is provided by the addition of fins or tubes. By these means, the height, size, and weight are held to desirable minimums. Special attention is given to sealing the transformers from the atmosphere. Likewise, careful attention is given to the external finish and fittings to assure reliable service for many years of exposure to the elements. External connectors are good for either aluminum or copper conductors.

The conventional pole type (Fig. 10-40) consists of core and coils securely mounted in an oil-filled tank, with the necessary terminals brought out through their appropriate bushings. The high-voltage bushings may be two in number, but one bushing plus a ground terminal on the tank wall connected to the ground end of the high-voltage winding for use on multiple-grounded circuits is the most common usage. The conventional type includes just the basic transformer structure without any protective equipment. The desired overvoltage, overload, and short-circuit protection is obtained by using lightning arresters and primary fuse cutouts separately mounted on the pole or crossarm closely adjacent to the transformer. The primary fuse cutout provides a means of visually detecting blown fuses on the system primary and also serves to remove the transformer from the high-voltage line, either manually when desired or automatically in the event of an internal coil failure.

The self-protected transformer (Fig. 10-41*a*) has an internally mounted, thermally controlled secondary circuit breaker for overload and short-circuit protection; an internally mounted protective link in series with the high-voltage winding to disconnect the transformer from the line in the event of an internal coil failure; and a lightning arrester or arresters integrally mounted on the outside of the tank for overvoltage protection. On most of these transformers, except some 5-kVA ratings, the circuit breaker operates a signal light when a predetermined winding temperature has been reached, as a warning before tripping. If the signal is unheeded and the breaker trips, the breaker may be reset and the load restored by an external handle. Usually, this can be accomplished with the normal breaker setting. If, however, the load has been a long-sustained one which has allowed the oil to reach a high temperature, the breaker may soon

FIGURE 10-40 Conventional pole-type distribution transformer.

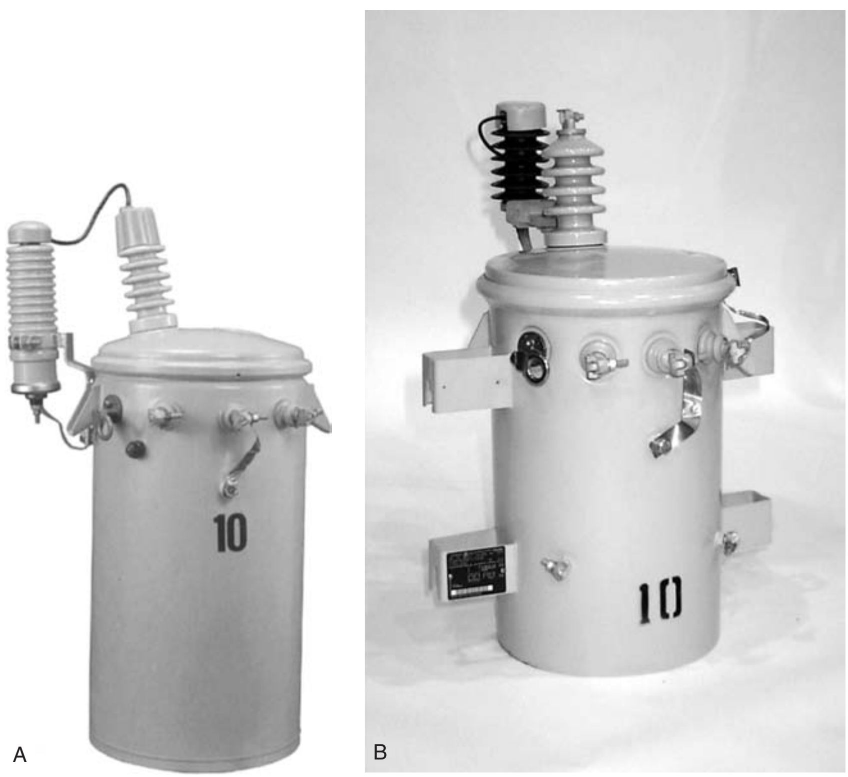


FIGURE 10-41 (*a*) Self-protected pole-type distribution transformer; (*b*) RUS design self-protected pole-type distribution transformer for rural electric cooperatives.

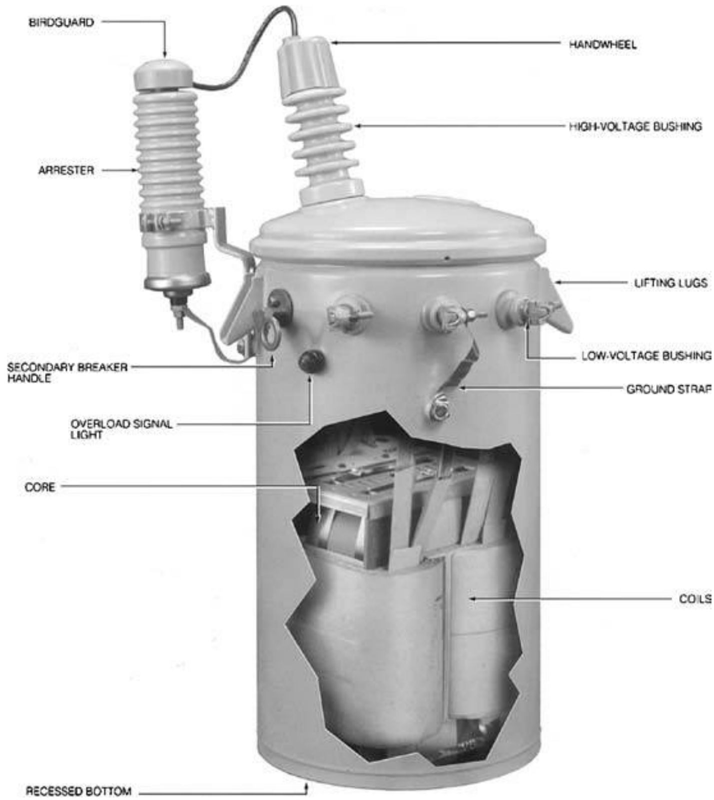


FIGURE 10-42 Cutaway view of a completely self-protected pole-type distribution transformer.

trip again, or it may be impossible to reset so that it will remain closed. In such cases, the trip temperature may be set up by an auxiliary external control handle to allow reclosing of the breaker for the emergency until a larger transformer can be installed.

Figure 10-41*b* shows the rural cooperative style. Figure 10-42 show a cutway view of the self-protected design.

3-Phase self-protected transformers are similar to the single-phase units except that a 3-pole circuit breaker is used. The breaker is arranged to open all 3 poles in case of a serious overload or fault on one of the phases (Fig. 10-43).

The self-protected transformer for secondary banking is another variation. Such transformers are provided with the two secondary breakers to sectionalize the low-voltage circuits, confining the outage to just the faulted or overloaded section, leaving the entire transformer capacity available for supplying the remaining sections. These are also made for single- and 3-phase units.

"Station-type" distribution transformers are normally rated 250, 333, or 500 kVA. A "pole/station type" distribution transformer is used. For distribution to low-voltage ac networks in areas of high-load density, network transformers are available in even higher ratings.

Losses and Characteristics. For the pole-type ratings 100 kVA and smaller, full-load efficiencies range from 97% to 99%, and impedance is generally less than 2%.

Recent trends have been toward further reduction of no-load loss, exciting current, and sound level. Replacement of much of the round and rectangular conductors in the secondary winding with full-width strip has resulted in more compact windings with greatly increased mechanical strength.

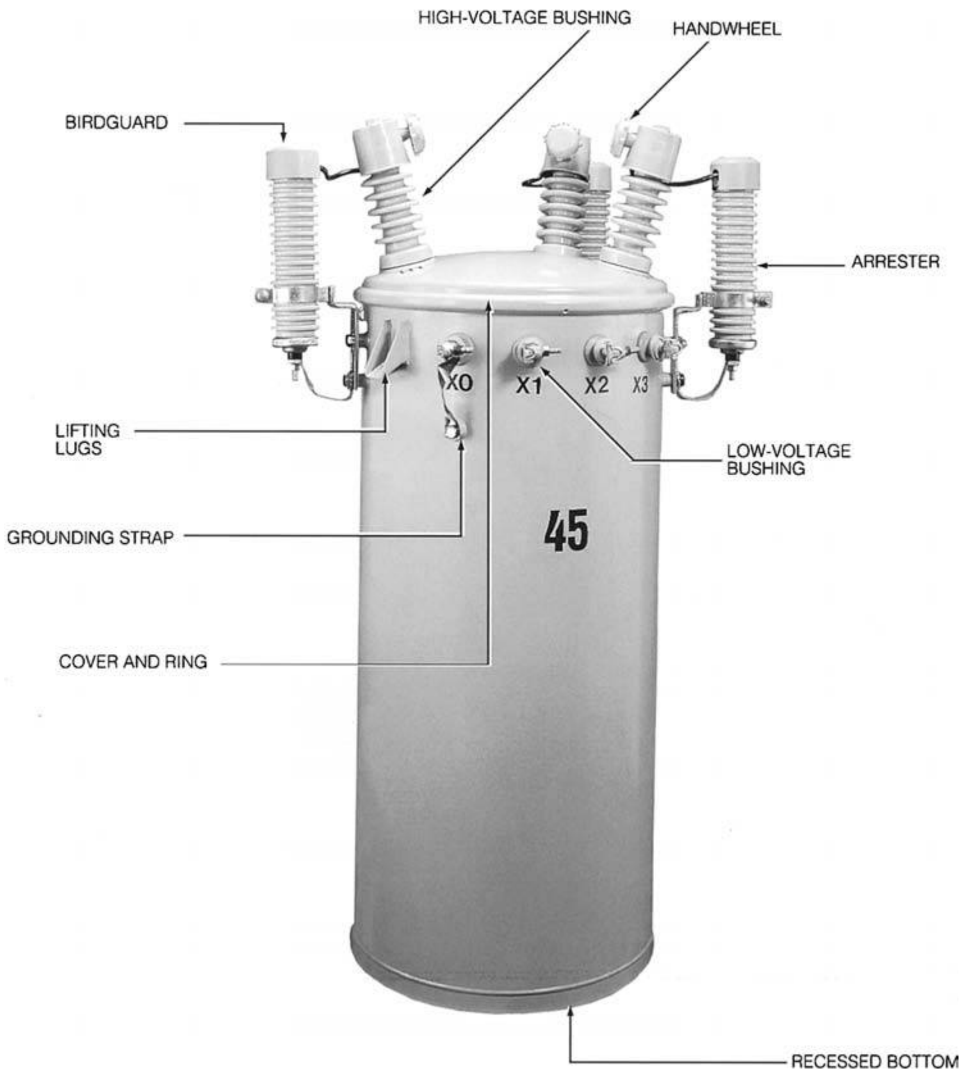


FIGURE 10-43 Three-phase, pole-type distribution transformer.

More effective utilization of distribution transformer investment is being made possible through transformer-load-management programs.

Transformers for Underground Distribution Systems. Since more distribution circuits are being put underground, transformers have been especially developed to be used with such systems. The most widely used type is the pad-mounted transformer, so called because it is designed to mount on a concrete-surface slab or pad. A typical transformer is shown in Fig. 10-44. The essential differences from the pole-type transformers are only in the mechanical arrangement:

1. A rectangular case in two compartments.
2. One compartment containing the conventional core-coil assembly.

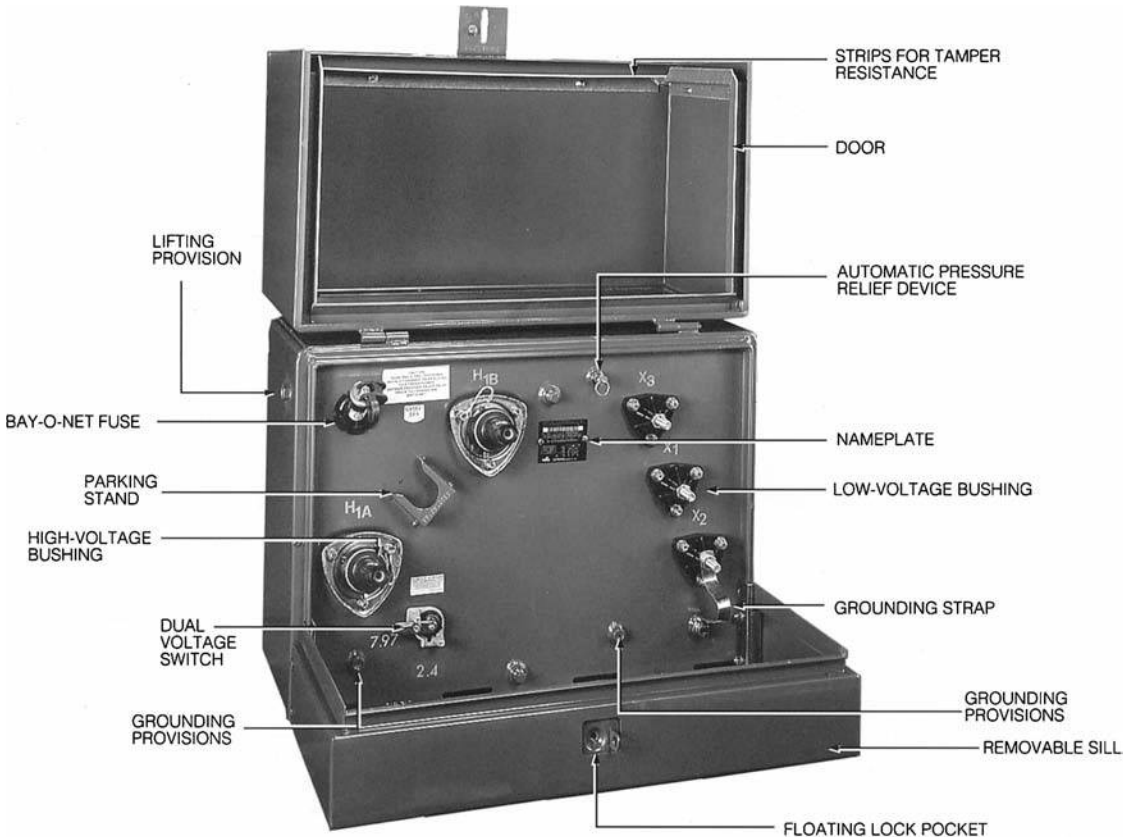


FIGURE 10-44 Single-phase, pad-mounted transformer.

3. A second compartment for cable termination and connection. The primary cable conductors are connected by plug-in connectors suitable for load make or break. The secondary conductors usually bolt to bushing terminals.
4. Fuses of various sorts provided for by a fuse holder placed in a well in the side of the tank so that the fuse holder can be drawn out.

Another transformer arrangement is designed to operate in a subterranean vault. This looks more like a pole-type transformer but is usually made with a corrosion-resistant steel tank, plug-in primary connectors, and a temperature rise in free air of only 55°C to allow for the higher ambient temperature which may actually exist in a vault.

Ferroresonance in Distribution Transformers. *Ferroresonance* is the name given to the phenomenon where the exciting reactance of the transformer can become nearly equal to the capacitive reactance of the line to ground, forming a resonant circuit. Such a resonant circuit can distort the normal line impedance to ground so that one line of a 3-phase circuit can rise to a destructive voltage.

Such a phenomenon practically never occurs in a normal circuit configuration with the transformers loaded, but it can exist under a combination of the following circumstances which usually occur only during switching of a 3-phase bank or blowing of a fuse in one line:

1. System neutral grounded, ungrounded transformer neutral
2. No load on the transformer
3. Relatively large capacitance line-to-ground such as may exist in cable circuits (underground distribution) or very long overhead lines (although ferroresonance can be and has been corrected by adding still more capacitance which presumably throws the combination out of resonance again)

Although ferroresonance has been studied at some length, it still does not seem possible to reliably predict its occurrence. Experience indicates that it is possible to prevent ferroresonance during switching on a transformer bank if all three transformers are resistance-loaded to 15% or more of their rating, or if special switches are used to assure that the three lines close simultaneously.

10.1.20 Furnace Transformers

Furnace transformers supply power to electric furnaces of the induction, resistance, open-arc, and submerged-arc types. The secondary voltage are low, occasionally less than 100 V, but generally several hundred volts. Sizes range from a few kVA to over 150 MVA, with secondary currents over 100,000 A. High currents are obtained by parallel connection of many winding sections. Current is collected by internal bus bars and brought through the transformer cover by the bus bars or by high-current bushings.

The power input to the furnace is controlled by adjusting the output voltage of the furnace transformer. Optimum performance of the furnace may require adjustment of the secondary voltage over a range of 3:1 or more. This may by accomplished by a regulating transformer between the high-voltage power source and a fixed-ratio furnace transformer. More frequently, regulation is obtained by taps in the high-voltage winding. In addition to taps in the high-voltage winding, a delta-Y switch in the high-voltage winding is often used to extend the range of voltage by an additional ratio of 1.73.

Motor-operated off-load tap changers are usual, but occasionally on-load tap-changing equipment is justified by the saving in melt time and reduced breaker maintenance. The load-tap-changing duty is more severe than on the usual power transformer, not only with respect to frequency of operation but also because of the extreme range, which results in large kVA increments per tap.

Circuit reactance furnishes current stability for ac arc furnaces. In the larger sizes, the inherent impedance of the transformer and its associated secondary conductors is sufficient for adequate stability. This is not generally true for smaller arc furnaces. Consequently, it is customary in furnace transformers rated 7500 kVA and below to include a reactor in the tank with the transformer. This reactor is connected in the high-voltage circuit and is furnished with taps to permit adjusting the total reactance to that required to maintain arc stability under the existing service conditions.

10.1.21 Grounding Transformers

A grounding transformer is intended primarily to provide a neutral point for grounding purposes. It may be a two-winding unit with a delta-connected secondary winding and a Y-connected primary winding which provides the neutral for grounding purposes, or it may be a single-winding 3-phase autotransformer with windings in interconnected Y or zigzag. With the latter, the windings consist of six equal parts, each designed for one-third the line-to-line voltage; two of these parts are placed on each leg and connected as in Fig. 10-45. In the case of a ground fault on any line, the ground current flows equally in the three legs of the autotransformer, and the interconnection offers the minimum impedance to the flow of the single-phase fault current.

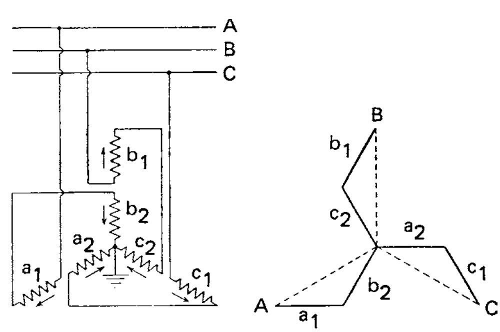


FIGURE 10-45 Grounding autotransformer with interconnected Y or zigzag windings.

10.1.22 Instrument Transformers

Functions. Instrument transformers are used to step transmission or distribution line voltages and currents down to levels that can be safely handled by metering and control devices. They are used to transform the primary voltage or current to values suitable for ratings of instruments, meters, relays, and other metering or control equipment. In a sense, they isolate the equipment from the lines. The normal secondary ratings are 5 A for current transformers and 115 or 120 V for voltage transformers. However, other current and voltage ratings are used depending on the application.

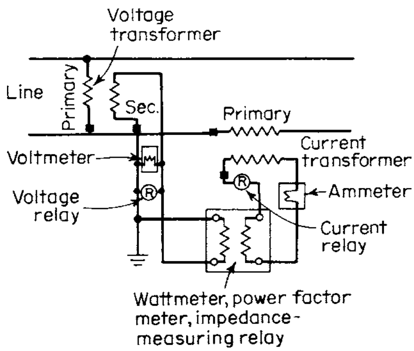


FIGURE 10-46 Voltage and current transformers as commonly connected to isolate meters, relays, and other equipment, and to transform current and voltage to convenient values.

To provide better protection, the secondary circuit should be grounded at one point. Metal cases should also be grounded.

The primary winding of a current transformer is connected in series with the load for which the current is to be measured or controlled. The primary winding of a voltage transformer is connected in parallel with the load for which the voltage is to be measured or controlled (Fig. 10-46). The secondary windings provide a current or voltage that is substantially proportional to the primary values for the operation of measuring instruments and control devices.

Polarity. When instrument transformers are used with measuring or control devices that respond only to the magnitude of the current or voltage, the direction of the current flow does not affect the response and the connections to the secondary terminals can be reversed without affecting the operation of the devices. When they are used with measuring or control devices that respond to the interaction of two or more currents, the correct operation of the devices depends on the relative phase positions of the currents, in addition to the magnitudes. To show the relative instantaneous directions of current flow, one primary and one secondary terminal are identified with a distinctive polarity marker, these indicate that at the instant when the primary current is flowing into the marked primary terminal the secondary current is flowing out of the marked secondary terminal (see Fig. 10-47).

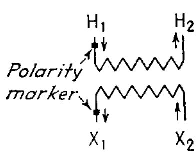


FIGURE 10-47 Polarity definition. Arrows indicate the instantaneous relative direction of currents in the windings.

Errors in Current Transformers. There are two types of errors that affect the accuracy of the measurements made with current transformers. The ratio-correction factor is the true ratio of the primary to the secondary current, divided by the nameplate ratio

$$F_{CR} = \frac{R_{CT}}{R_{CN}} \tag{10-66}$$

where F_{CR} = ratio correction factor of current transformer, R_{CT} = true ratio [(primary current)/(secondary current)], and R_{CN} = nameplate ratio [(primary current)/(secondary current)] of current transformer.

The phase-angle error is the angle of lead of the current leaving the marked secondary terminal over the current entering the marked primary terminal.

Relative Importance of Ratio and Phase Angle Errors. The ratio correction factor of current transformers affects the magnitude of the outputs. Therefore, they are important for meters that measure power for revenue purposes or other applications where the magnitude of the current is important. For example, a ratio correction factor of 1.010 indicates that the secondary current is lower than the correct value by 1% and that all measuring or control devices connected in the secondary circuit will have 1% less current than the primary current divided by the marked ratio.

The phase-angle error does not affect current-actuated devices such as ammeters or overcurrent relays, but the accuracy or operation of devices that respond to the products, the sums, or the differences of currents are affected by the phase-angle error. A wattmeter is a device that responds to the product of the voltage applied to the potential terminals, the current through the current coils, and the power factor which is the cosine of the angle between the voltage and current. The phase-angle error becomes very important if the power factor of the load being measured, such as measurement of the load losses in large power transformers, is small because the magnitude of the phase-angle error is significant compared to the angle between the voltage and the current in the devices being measured. If the current is supplied from the secondary of a current transformer with unity nameplate ratio, unity ratio correction factor, a phase-angle error of β, and primary current lagging the voltage, the wattmeter will not indicate the true watts, $EI \cos\theta$, but will indicate $EI \cos(\theta - \beta)$. If the sign of β is plus, the $\cos(\theta - \beta)$ will be larger than the $\cos\theta$ and the wattmeter will read high (see Fig. 10-48). If the sign of β is minus, the wattmeter will read low. To obtain the true watts, the apparent watts should be multiplied by the phase-angle correction factor K_β, which is dependent on both the phase-angle error of the transformer and the power factor of the load, as follows:

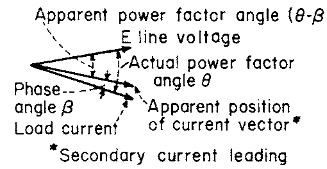


FIGURE 10-48 Effect of positive phase angle in increasing the apparent power factor.

$$K_\beta = \frac{\cos\theta}{\cos(\theta - \beta)} = \frac{1}{\cos\beta + \sin\beta\tan\theta} \tag{10-67}$$

where K_β = phase-angle correction factor of current transformer, β = angle of lead of secondary current over primary current, and θ = angle of lag of load current behind load voltage.

Summarizing the Effect of Current-Transformer Errors. Ratio correction factors: Above 1.000 the secondary current is low, and below 1.000 the secondary current is high. Phase-angle errors with lagging load current: Positive phase-angle errors cause wattmeter readings to be high if the load-current power-factor angle is greater than the phase-angle error, and negative phase-angle errors cause the wattmeter readings to be low.

In practical metering problems β will be less than 30 min and K_β can be written

$$K_\beta = 1 - \sin\beta\tan\theta = 1 - \tan\theta \qquad \text{where } \beta \text{ is in radians}$$
$$K_\beta = 1 - \beta\tan\theta/3438 \qquad \text{where } \beta \text{ is in minutes} \tag{10-68}$$

The uncertainty in knowledge of the exact values of β and θ represents a greater error than will result from this simplification of Eq. (10-67).

Transformer Correction Factor. The transformer correction factor to be applied to the reading of the wattmeter for both the ratio and phase-angle errors is given by Eq. (10-69)

$$F_T = F_{CR}K_\beta \tag{10-69}$$

where F_T = transformer correction factor.

If F_{CR} and K_β are both between 0.985 and 1.015, Eq. (10-70) may be used with an error under ± 0.0003:

$$F_T = F_{CR} + K_\beta - 1.000 \tag{10-70}$$

Classification of Errors. ANSI C57.13 classifies current transformers as to accuracy by a method which limits the total error in a wattmeter or watthour-meter reading resulting from the combination of ratio and phase-angle errors, over a range of power factors of the metered load of 0.6 to 1.0. The most accurate classification (usually specified for use with watthour meters for billing metering) is 0.3, which means that the total error at the meter caused by the current transformer will not exceed 0.3% at rated current (or at maximum continuous current), 0.6% at 10% rated current. ANSI C57.13 also recognizes 0.6 and 1.2 accuracy classes. The accuracy class is specified in connection with one or more of the standard burdens (see the next paragaph). For example, "0.3 B-0.2" describes a transformer of 0.3 accuracy class when loaded with a B-0.2 burden on the secondary terminals.

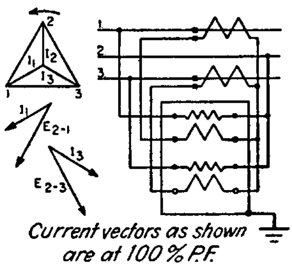


FIGURE 10-49 Metering connections for 3-phase, 3-wire lines with two watthour-meter elements.

Standard Burdens. ANSI C 57.13 also recognizes a number of standard values of secondary impedance loading (called "burden" in instrument-transformer parlance) for use in describing current transformer performance. The burdens are designated as B0.1, B0.2, B0.5, B0.9, B1.8 to mean impedances of 0.1, 0.2, etc., ohms at 0.9 power factor, and also B1.0, B2.0, B4.0, and B8.0 for corresponding impedances at 0.5 power factor. (These standard burden impedances are only applicable if rated secondary current is 5 A.)

Effect of Phase-Angle Errors on Metering 3-Phase 3-Wire Power with a Two-Stator Watt-hour Meter. With the meter connected as indicated in Fig. 10-49 with balanced load, if the marked ratio and ratio correction factor are both 1.000 and the phase-angle error β is the same for both current transformers

$$K_\beta = \frac{E_1 \cos(\theta + 30) + E_1 \cos(\theta - 30)}{[E_1 \cos(\theta + 30 - \beta) + E_1 \cos(\theta - 30 - \beta)]} = \frac{1}{\cos\beta + \sin\beta \tan\theta} \qquad (10\text{-}71)$$

This is identical with Eq. (10-67) for a single-phase measurement.

Causes of Errors. The errors in a current transformer are due to the energy required to produce the core flux, which induces the secondary winding voltage that supplies the current through the secondary circuit. The total ampere-turns available to provide secondary current are vectorially equal to the primary ampere-turns minus the ampere-turns required to produce the core flux.

A change in secondary burden alters the flux required in the core and changes the core-exciting ampere-turns; leakage flux entering the core changes the magnetic characteristics of the core and affects the core-exciting ampere-turns.

Calculation of Ratio and Phase-Angle Errors. Caculation of ratio and phase angle in current transformers can, in theory, be done simply by first calculating the exciting ampere-turns required to magnetize the core and subtracting them from the primary ampere-turns to get the secondary ampere-turns (vectorial subtraction). The difficulty in calculation is that in many transformers the leakage flux which flows in only part of the core is as large as the working flux and its effect on exciting ampere-turns is most difficult to calculate (Wentz 1941). For exact determination of ratio and phase angle, therefore, actual measurement of ratio and phase angle as described in ANSI C57.13 is necessary.

Approximate calculation of ratio for relaying service is more practical and in fact often necessary because transformers supplying relays must often operate at currents up to 20 times normal or even higher, and exact measurements are extremely difficult and expensive. Accordingly, ANSI C57.13 recognizes an adequate ratio-error calculation method for current transformers in which the leakage flux which enters the core has only a negligible effect on performance. This calculation method is described in the Test Methods Section of C57.13. The ratio error calculated by this method may in fact be greater than the actual ratio error, but it is certain that it will not be less, a conservative result.

Generally, the leakage flux in a current transformer can be said to be negligible only in a current transformer consisting of a toroidal core with the secondary winding fairly well distributed around the core, for example, an assembly intended to be placed over a bushing in a circuit breaker or transformer and then only if the return conductor is at a distance at least as great as in typical circuit breakers. If this type of transformer is used with a primary coil or even a single primary turn placed around one side of the core, it is doubtful that the leakage flux will be negligible.

Transformers with negligible leakage flux (principally bushing type) are designated in C57.13 as Class C transformers, meaning that it is safe to calculate the ratio error.

Classification of Current Transformers for Relaying Service. Transformers for which calculation of ratio will give conservative results are designated as Class C; all other transformers are designated as Class T, meaning that their performance must be determined by test. The performance of both types of transformers is then classified according to the voltage the secondary can deliver to the burden at 20 times rated secondary current without exceeding 10% ratio error. The established standard secondary terminal voltages are 10, 20, 50, 100, 200, 400, and 800, corresponding to the voltage required by ANSI standard burdens at 20 times normal current (100 A). A current transformer is classified by the letter C or T plus the standard voltage, such as C200 or T400.

Short-Time Current Limits. Current transformers may have to carry very large currents in the event of a short circuit of the system, especially when a low-rated branch circuit is supplied from a large system with a high fault-current capability. The large currents in the winding have two principal effects:

1. The primary winding is repelled from the secondary by the electromagnetic forces, which are proportional to the square of the current.

2. The windings heat very rapidly at a rate nearly proportional to the square of the current.

In order to apply current transformers properly, it is necessary to give them ratings:

1. Mechanical short-time rating, the current, usually stated in terms of times normal, which the transformer can withstand mechanically even if the current is initially fully offset.

2. Thermal short-time rating, the current which will not heat the winding to more than 250°C for copper conductors or 200°C for EC aluminum conductors. This limit is not to be demonstrated by test but is calculated on the conservative assumption that all the heat generated by the current is stored in the conductor.

C57.13 may be consulted for additional detail and for the test method to be used to demonstrate the mechanical limit.

Types of Construction. The types of current transformers are the wound type which consists of primary and secondary windings completely installed and permanently assembled on the magnetic circuit; the bar type, which is similar to the wound type except that the primary is a single straight and fixed turn; the window type, which has a secondary winding completely insulated and permanently assembled on the magnetic circuit and a window through which a conductor can be passed to provide a primary winding; and the bushing type, which is a special window type designed to fit over apparatus bushings, with the conductor through the bushing as the primary winding.

Current transformers are classified in accordance with the major insulation used as dry-type, compound-filled, molded, or liquid-immersed.

Safety Precautions. The secondary winding should always be short-circuited before disconnecting the burden. If the secondary circuit is open, with primary current flowing, all the primary ampere-turns are magnetizing ampere-turns and usually will produce an excessively high secondary voltage across the open circuit. All instrument-transformer secondary circuits should be connected to ground; when instrument-transformer secondaries are interconnected, only one point should be grounded. If the secondary circuit is not grounded, the secondary becomes, in effect, the middle plate of a capacitor, with the high-voltage winding and ground acting as the other two plates.

Errors in Voltage Transformers. There are two types of errors that affect the accuracy of the measurements made with voltage transformers. The ratio error is the difference between the true ratio of the primary to secondary voltage and the ratio that is marked on the nameplate. The phase-angle error is the difference in the phase position of the voltage applied to the secondary burden and the voltage applied to the primary winding.

The ratio error is expressed as a ratio correction factor by which the secondary-voltage value should be multiplied to obtain a secondary voltage that is directly proportional to the primary voltage

$$F_{PR} = \frac{R_{PT}}{R_{PH}} \tag{10-72}$$

where F_{PR} = ratio correction factor of the voltage transformer, R_{PT} = true ratio (primary/secondary) of the voltage transformer, and R_{PN} = nameplate ratio (primary/secondary) of the voltage transformer.

The phase-angle error is designated by the symbol γ, is expressed in minutes, and is defined as positive when the voltage applied to the burden from the marked to the unmarked secondary terminal leads the voltage applied to the primary from the marked to the unmarked terminal.

Relative Importance of Ratio and Phase-Angle Errors. The effect of the ratio and phase-angle errors of voltage transformers is the same for current transformers, except that with a lagging power-factor load a positive voltage transformer phase-angle error will cause the wattmeter to read low. To obtain the true watts, the apparent watts should be multiplied by the phase-angle correction factor K_γ:

$$K_\gamma = \frac{\cos\theta}{\cos(\theta + \gamma)} = \frac{1}{\cos\gamma - \sin\gamma\tan\theta} \tag{10-73}$$

where K_γ = phase-angle correction factor of the voltage transformer, θ = angle of lag of load current behind load voltage, and γ = angle of lead of secondary voltage over primary voltage.

Classification of Errors. The errors in a voltage transformer change with the current required by the burden connected across the secondary terminals, the frequency, and the magnitude of the secondary voltage. ANSI C57.13 classifies voltage transformers as to accuracy by a method essentially the same as that for current transformers, the principal difference being that the limits of error apply over the range of rated voltage from 90% to 110% and from zero burden up to the burden at which the rating is given. The standard burdens for rating purposes are given in Table 10-8. The complete ANSI accuracy classification of a voltage transformer must include the secondary burden, such as 0.3X or 0.6Z.

Voltage transformers are made for all the standard rated circuit voltages. They are usually dry-type or molded for voltages below 23 kV and liquid-filled for the higher voltages.

Other Error Calculations and Corrections. The measurement of power by using a wattmeter and both current and voltage transformers requires a correction for the ratio and phase-angle errors of both of the instrument transformers and a correction for the phase angle α in minutes of the potential circuit of the wattmeter. Figure 10-50 shows the relative phase positions of the primary and secondary

TABLE 10-8 Standard Burdens for Voltage Transformers

From Table 13-11.110, ANSI C57.13

Burden designation	Secondary voltamperes*	Burden power factor
W	12.5	0.10
X	25	0.70
M	35	0.20
Y	75	0.85
Z	200	0.85
ZZ	400	0.85

*At 120 or 69.3 secondary volts, if rated secondary voltage is from 90% to 110% of 120 or 69.3 V.

voltages and currents of the instrument transformers if both β and γ are positive. If the potential circuit of the wattmeter is inductive by a small amount, the current in this circuit will lag the voltage and the phase angle will have the same effect as a negative voltage-transformer phase angle. The total phase-angle correction factor is

$$K_S = \frac{\cos(\theta + \beta - \gamma + \alpha)}{\cos\theta} \tag{10-74}$$

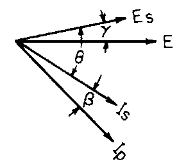


FIGURE 10-50 Vector relations in current and voltage transformers, where p = 5 primary and s = 5 secondary.

where K_S = total phase-angle correction factor, α = angle of lag of potential circuit of wattmeter, β = angle of lead of secondary current over primary current, γ = angle of lead of secondary voltage over primary voltage, and θ = angle of lag of secondary current behind secondary voltage.

The true watts, with all corrections included, is

$$P = WR_{CN}R_{PN}F_{CR}F_{PR}K_S \tag{10-75}$$

where F_{CR} = ratio-correction factor of current transformer, F_{PR} = ratio-correction factor of voltage transformer, P = watts drawn by load, R_{CN} = nameplate ratio of current transformer, R_{PN} = nameplate ratio of voltage transformer, and W = watts reading of wattmeter.

In watt-hour meters, the phase angle corresponding to α is corrected for in the meter by the lag adjustment or by making an overall calibration and adjustment for all errors, including the ratio and phase-angle errors of the instrument transformers.

Figures 10-51 and 10-52 show typical ratio-correction-factor and phase-angle curves for a current and a voltage transformer.

Consider the following example in which a load is measured with 50/5-A current transformer, with a 12.5-VA 90% power-factor burden characteristic curve according to Fig. 10-51 with 2300:115 V voltage transformer rated 200 VA with a 75-VA 85% power-factor burden characteristic curve according to Fig. 10-52 and with a lag angle of 6 min in the wattmeter potential circuit. What is the load true watts when the instruments read 500 W, 115 V, and 5 A, calibration errors of instruments being neglected? From Fig. 10-51

$$F_{CR} = 0.9997 \qquad \beta = -2'$$

From Fig. 10-52

$$F_{PR} = 0.9984 \qquad \gamma = +1$$

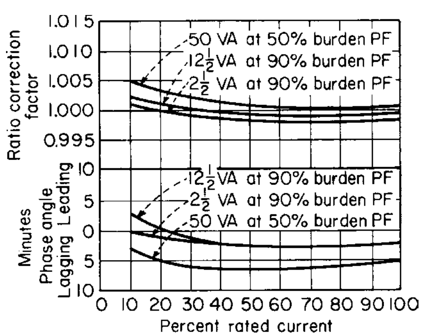


FIGURE 10-51 Typical curves of the ratio correction of phase angle for a current transformer.

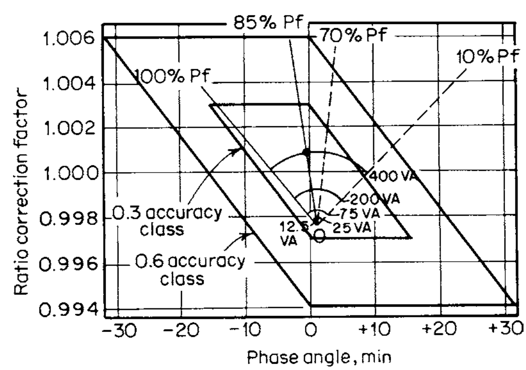


FIGURE 10-52 Typical curves of ratio-correction factors and of phase angle for a voltage transformer.

Stated above

$$\alpha = +6'$$

$$\theta = \cos^{-1}\frac{500}{5 \times 115} = \cos^{-1} 0.86957 = 29°35.47'$$

From Eq. (10-74)

$$K_S = \frac{\cos(29°35.47' - 2' - 1' + 6')}{\cos(29°35.47')} = \frac{0.86914}{0.86957} = 0.9995$$

From Eq. (10-75)

$$P = 500 \times 10 \times 20 \times 0.9997 \times 0.9984 \times 0.9995 = 99760$$

Ferroresonance in voltage transformers can occur when a voltage transformer is connected line to ground on an ungrounded system if the capacitive reactance, line to ground, can equal the exciting reactance of the voltage transformer, a condition which typically occurs at considerable overvoltage. A parallel resonant circuit of very high impedance is formed, raising the line-to-ground voltage considerably above normal; if the voltage is high enough, the voltage transformer will fail due to excessive exciting current or high-voltage stresses.

Measures taken to reduce the incidence of ferroresonance include operation at low flux densities, resistance loading of the secondaries, and insertion of resistance in the connection of the voltage transformer primary neutral to ground.

Transient performance of current transformers with initially offset primary current has been given a great deal of study in recent years. It has been realized for many years that a transformer has difficulty in transforming the dc component of an offset current. No one method of calculation has been generally acceptable but study of the references will help one to arrive at a solution of most practical problems.

10.2 CIRCUIT BREAKERS

By DAVID S. JOHNSON, JEFFREY H. NELSON and TED W. OLSEN

Definitions of terms used in this section can be found in the IEEE standards and application guides referenced in this section and/or in the *IEEE Authoritative Dictionary of Terms*.

10.2.1 Fundamentals

Definitions. *Circuit breakers* are mechanical switching devices capable of making and breaking currents under either normal or specified abnormal (short circuit) conditions on the power system. Though circuit breakers are primarily defined by their protective capabilities and ratings under abnormal short circuit conditions, they also perform switching duties under a myriad of other system conditions, each of which has its own set of switching stresses.

Circuit breakers are rated primarily by power frequency voltage, insulation levels (BIL, switching impulse, hi-pot voltage), continuous current, short-circuit current, and interrupting time. Reference is made to IEEE C37.100[1] for definitions of ratings subjects, and to IEEE C37.04[2] and IEEE C37.06[3] for values of ratings typically applied to circuit breakers.

Circuit breakers employ a variety of media for high voltage insulation and/or current interruption. The type of media employed in a specific design is often designated as a prefix in the naming of the circuit breaker, for example, vacuum circuit breaker, or sulfur hexafluoride (SF_6) gas circuit breaker.

Circuit breakers are often categorized as being of either "dead tank" or "live tank" design. In the dead tank case, the interrupting contact system is enclosed in a grounded tank, typically surrounded by an insulating fluid (oil) or gas (SF_6) (see Figs. 10-53 and 10-54 for examples of dead tank circuit breakers).

The electrical current enters the tank through high voltage entrance bushings (Fig. 10-55), passes through the contact system, and then exits through another high voltage entrance bushing. In the live tank case, the interrupting contact system is supported by insulators at some height above ground potential, but is not contained within a grounded tank system (see Figs. 10-56 and 10-57 for examples of live tank circuit breakers). There is no grounded tank or enclosure surrounding the live parts. Dead tank design allows the placement of current transformers, which are necessary for protective relaying input signals, around the high voltage entrance bushing. Live tank design offers no location to place current transformers, and therefore must be independently placed adjacent to the circuit breaker.

History of Development. In the early days of electrification (1890), switches were of the hand-operated, knife-blade type.

Air Switches. With increasing current and voltages, spring-action driving mechanisms were developed to reduce contact burning by faster-opening operation. Later, main contacts were fitted with arcing contacts of special material and shape, which opened after and closed before the main contacts. Further improvements of the air

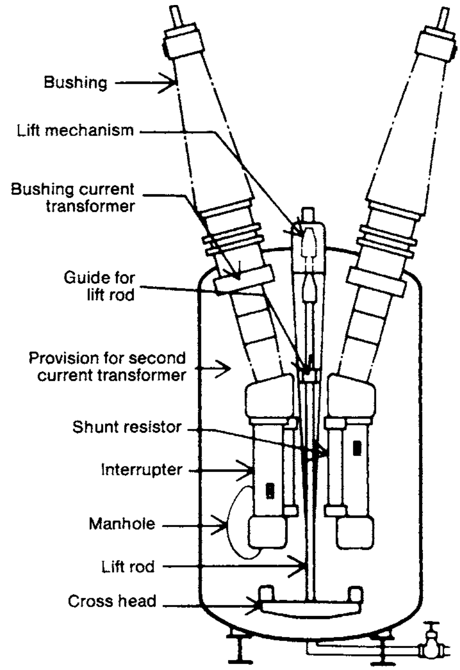


FIGURE 10-53 Outline of a dead-tank 161-kV outdoor oil circuit breaker.

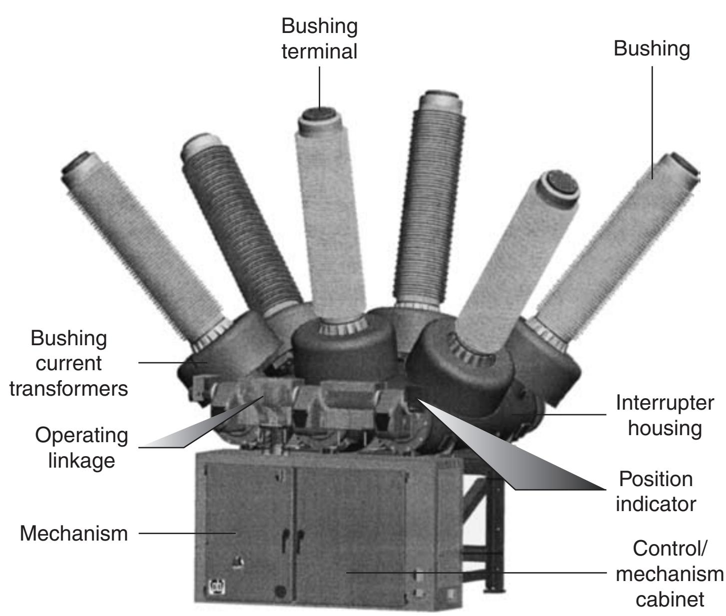


FIGURE 10-54 Dead-tank SF_6 circuit breaker rated 245 kV, 63 kA.

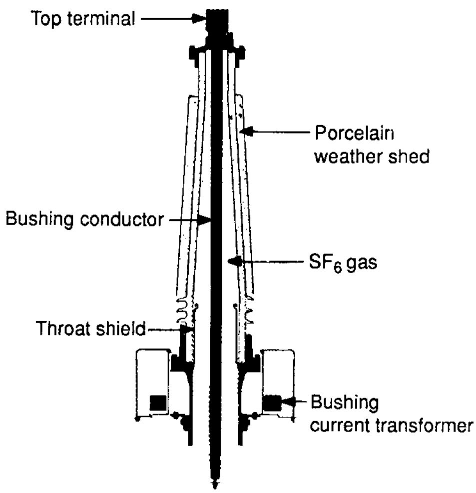


FIGURE 10-55 Gas-filled bushing.

switch were the brush-type contact with a wiping and cleaning function, the insulating barriers leading to arc chutes, and blowout coils with excellent arc-extinguishing properties. These features, as well as the horn gap contact, are still in use in low-voltage ac and dc breakers.

Oil Circuit Breaker. Around 1900, in order to cope with the new requirement for "interrupting capacity," ac switches were immersed in a tank of oil. Oil is very effective in quenching the arc and establishing the dielectric strength of an open break after current zero. Deion grids, oil-blast features, pressure-tight joints and vents, new operating mechanisms, and multiple interrupters were introduced over several decades to make the oil circuit breaker a reliable apparatus for system voltages up to 362 kV.

Minimum-Oil Circuit Breaker. These breakers were developed after 1930, used mainly in Europe, and make the special low-oil-volume interrupting chambers of extra-light weight. By means of current-dependent oil streams in different directions and supported by oil injection, the arc is cooled and extinguished effectively. The interrupters are mounted on porcelain or molded-resin supports, thus avoiding oil as an insulating medium to ground.

Standard transformer oil can be used for both oil and minimum-oil circuit breakers.

Air-Blast Circuit Breaker. Further increase of system voltages and generating capacities triggered the search for faster and stronger circuit breakers utilizing oiless arc interruption. After 1940, the air-blast circuit breaker was developed, making use of the good insulating and arc-quenching properties of dry and cleaned compressed air.

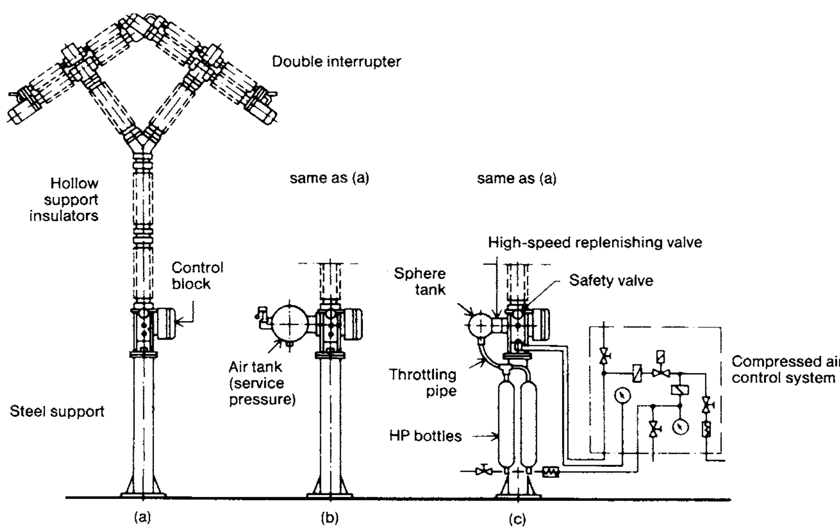


FIGURE 10-56 Modular setup of a 362-kV outdoor air-blast circuit breaker with two uprating steps: (*a*) without air tank, low braking capability; (*b*) with air tank, standard breaking capability; (*c*) with constant-pressure air supply and high breaking capability.

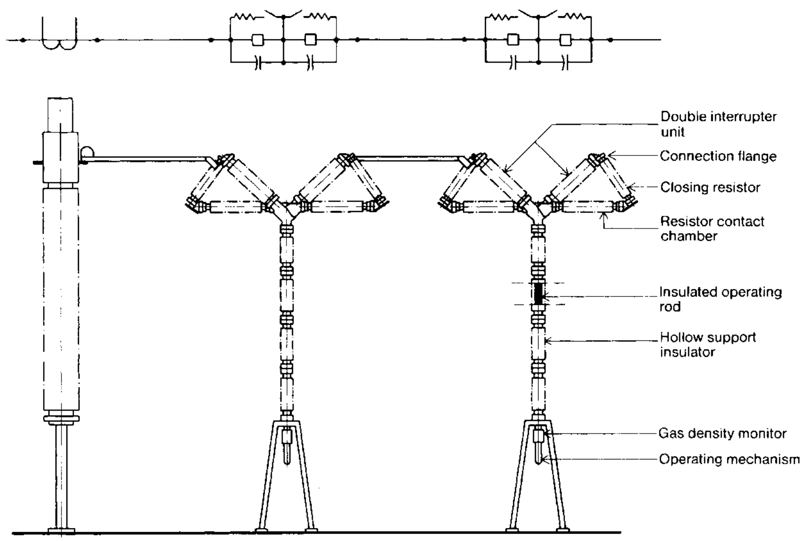


FIGURE 10-57 Live-tank SF_6 outdoor circuit breaker and current transformer arrangement, 800 kV, 3000 A, 40 kA.

Figure 10-58 shows a typical air-blast circuit breaker of modular design, installed in 1950.

Further development of the air-blast circuit breaker led to two-cycle interrupting time, extra-heavy interrupters, and the constant-pressure control system.

The magnetic air circuit breaker uses a combination of a strong magnetic field (coil or soft iron plates) with a special arc chute to lengthen and cool the arc until the system voltage cannot maintain the arc any longer. This interrupting principle was applied mainly in the distribution voltage range in metal-clad switchgear from the 1940s to the 1980s.

SF_6 Gas Circuit Breakers. SF_6 gas circuit breakers were first developed in the early 1950s by Westinghouse Corporation, following the discovery of the excellent arc quenching and insulating properties of SF_6 gas (see Fig. 10-59). Both live tank and dead tank designs were introduced from the late 1950s into the 1960s. SF_6 remains the dominant insulating and arc-quenching medium at higher voltages (72.5 kV and above) even today.

Dead tank SF_6 gas circuit breakers were incorporated into gas-insulated substations (GIS) up to 800 kV from the mid-1960s through the present. Gas insulated substations offer space savings and

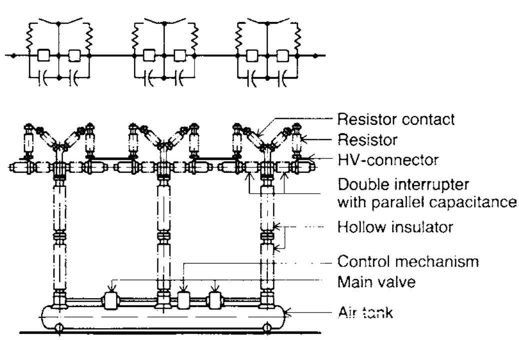


FIGURE 10-58 Outdoor air-blast circuit breaker 230 kV, 10009 A, 16 kA, 3-cycle interrupting time.

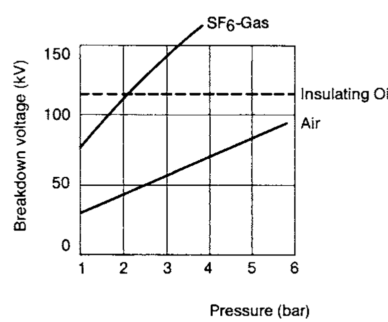


FIGURE 10-59 Breakdown voltage of oil, air, and SF_6 gas as a function of pressure at 38 mm ($1\frac{1}{2}$ in) electrode distance.

environmental advantages over conventional outdoor substations, using the reduced insulation gap requirements of SF_6 gas.

SF_6 gas circuit breakers were initially of the two-pressure type, in which high pressure gas for interruption is compressed and stored for later interrupting duty. Later designs employed the puffer principle, in which interrupting pressure is developed during the contact motion itself, and no high pressure gas is stored. The latest designs of SF_6 gas circuit breakers utilize the arc thermal energy itself to develop the interrupting pressure; these designs are referred to as *self-blast* or *thermal-assist* circuit breakers.

Vacuum Circuit Breakers. Vacuum interrupter technology is a relatively recent development, from the 1960s to 1970s. Vacuum circuit breakers employ a vacuum bottle, which includes the contact system. The contacts within the vacuum bottle are driven, through a bellows, by a low energy mechanism. The contact stroke of vacuum circuit breakers is short, which leads to very reliable, low mechanical-energy designs. Vacuum interrupters have now improved in their range of voltages, short-circuit currents, and continuous current capabilities, such that most medium voltage application requirements can be met. Vacuum circuit breaker designs have become the dominant technology for outdoor breakers and metal-clad switchgear applications at 38 kV and below.

Design Fundamentals. Circuit breaker designs typically consist of the following construction elements: (1) contact system or interrupter, at high voltage; (2) insulation between the contact system and ground potential (SF_6 gas, porcelain, molded resin); (3) operating mechanism and related control systems; and (4) an insulated link between the operating mechanism and the high voltage contact system. In addition, dead tank circuit breakers incorporate high voltage entrance bushings that carry current at high voltage through the circuit breaker tank shell. Dead tank designs may also include bushing type current transformers (BTCT), which are conveniently placed around the conductor of the high voltage entrance bushings.

Tripping facilities. Tripping facilities, including circuit breaker controls, and the mechanical tripping mechanism of the circuit breaker operating mechanism, are vital to ensure proper operation during all conditions.

Short Circuit Duty. The *short-circuit duty* is determined by the maximum short-circuit that the rotating machinery connected to the system at the time of short circuit can pass through the breaker to a point just beyond the breaker, at the instant the breaker contacts open. The short-circuit current is determined by the characteristics of synchronous and induction machines connected to the system at the time of the short circuit, the impedance between them and the point of short circuit, and the elapsed time between the starting of the short circuit and the parting of the breaker contacts.

In *calculating* short-circuit currents of high-voltage ac circuit, it is ordinarily sufficiently accurate to take into account only the reactance of the machines and circuits, whereas in low-voltage circuit resistance as well as reactance may enter into the calculation. In dc circuit, resistance only is ordinarily sufficient.

For first *approximations*, the reactance and typical time-decrement curves of the synchronous machines may be used. For close calculations, the actual reactances and time characteristics of the equipment should be used, and calculation made for single- as well as 3-phase faults. The “per unit” impedance system and the “internal voltage” method, using “symmetrical components,” are often used in more exact calculations. Programs are available for digital computer studies of system short-circuit currents, both balanced 3-phase and phase-to-ground.

The *interrupting capacity*, in kilovolt amperes, is the product of the phase-to-ground voltage, in kilovolts, of the circuit and the interrupting ability, in amperes, at stated intervals and for a specific number of operations. The current taken is the rms value existing during the first half-cycle of arc between contacts during the opening stroke.

Symmetrical Current Basis. It has become a widely adopted practice to determine the interrupting capability of circuit breakers in kiloamperes symmetrical. The rated short-circuit current in rms kiloamperes is referred to the rated maximum voltage in kilovolts.

The ratings structure and tables of ratings for ac high-voltage circuit breakers are found in IEEE C37.04[2] and IEEE C37.06[3].

The short-circuit current interrupting process is characterized first by an arc appearing between the breaker contacts. The arc contains a high conductivity plasma column originating from the high temperature and related gas ionization (in the case of gas-blast interrupters). Interruption will occur at current zero and in this case is first determined by successful cooling of the arc (through gas flow) to eliminate the ionized gas conductive path, and then the race to build up dielectric strength of the open contact gap faster than the rise of the power system recovery voltage.

Several specific problems are encountered during the interrupting process of gas-blast interrupters:

1. Arc plasma temperatures exceeding 20,000 K.
2. The turbulent supersonic flow of the quenching gas in a changing flow geometry with speeds ranging from a few hundred meters per second to several thousand meters per second.
3. The interrupter-moving system and its drive accelerates the moving masses in the few thousandths of a second to speeds as high as 10 m/s while simultaneously compressing the quenching gas.
4. The stress places on the network system by the current interruption and the recovery voltage.

The *interrupting principle* of an SF_6 puffer-type interrupter is sketched in Fig. 10-60. On opening, the fixed and moving contacts are pulled apart by the operating mechanism. Thus, the fault current is forced to flow along the arc plasma. The contact movement combined with the compression cylinder movement in the opposite direction compresses the quenching gas inside the cylinder. The quenching gas is consequently forced to flow through the contact system, and the insulated nozzle toward the exhaust. This intensive flow of quenching medium along the arc rapidly removes the energy converted within the arc plasma and transforms the path between the open contacts into an insulating gap.

DC Interruption. DC interruption (see Fig. 10-61) is basically different from ac interruption. After contact parting, the arc is lengthened and cooled and consequently the arc voltage is rising. The current will extinguish after the change in current *di*/*dt* becomes negative and the arc voltage rises above service voltage. DC circuit breakers must therefore operate fast, in order to allow the voltage to build up in a few milliseconds. The energy originating from the generator and inductance will have to be absorbed by the arc. Magnitude and duration of short-circuit current depend on the height of network inductance. With rising inductance, the short-circuit current will decrease, whereas the duration of the arc will increase.

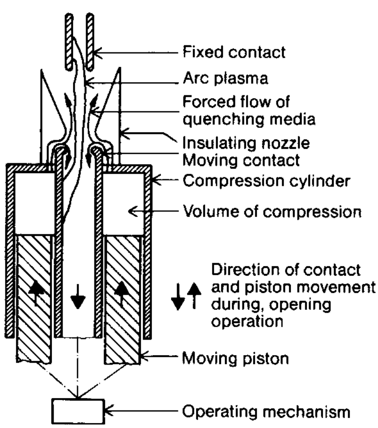


FIGURE 10-60 Principle of puffer-type arc interrupter.

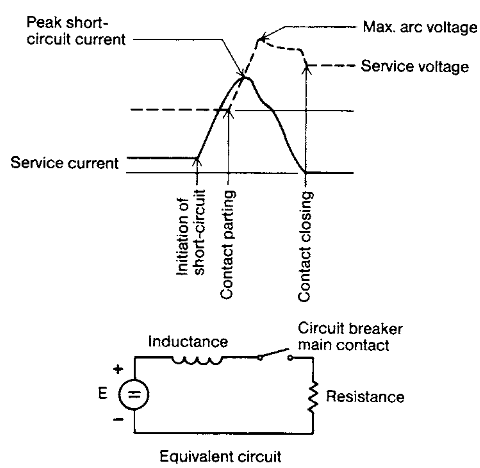


FIGURE 10-61 Direct-current interruption; typical shape of short-circuit current and recovery voltage.

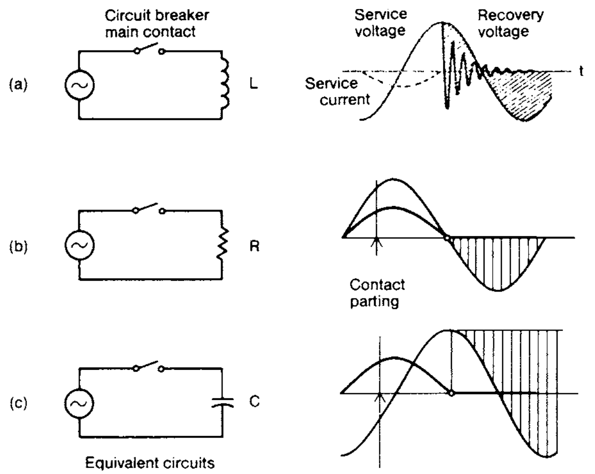


FIGURE 10-62 Typical shape of recovery voltage on interruption: (*a*) induction current; (*b*) resistive current; (*c*) capacitive current.

AC Interruption. AC interruption occurs at current zero. During the following half-cycle, the *recovery voltage* will build up across the circuit breaker main contacts. The typical appearance of recovery voltage will differ in inductive, resistive, and capacitive circuits (see Fig. 10-62). When opening an inductive circuit, the recovery voltage will rise suddenly at a high rate because current interruption occurs at the moment of system voltage peak. This case requires fast building of dielectric strength of the open contact gap.

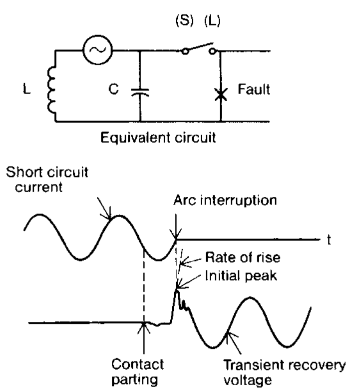


FIGURE 10-63 Alternating-current interruption; typical shape of short-circuit current and transient recovery voltage.

When interrupting resistive load, current and voltage pass through zero at about the same moment. The recovery voltage will therefore rise at a moderate rate and no particular problems are imposed on the circuit breaker.

At the moment of interruption of capacitive current, the capacitance is fully charged. The recovery voltage rises slowly during the first half cycle but continues to rise to a value twice the system voltage. This may lead to restrikes, undesired network oscillations, and overvoltages.

The waveforms of short-circuit current and transient recovery voltage in a simplified network system are shown in Fig. 10-63. At the moment of fault-current interruption the two sections—source side (*S*) and line side (*L*)—of the network are decoupled and oscillate independently about their driving voltage. The difference of these two transients appears across the open contacts of the breaker pole. The behavior of this transient recovery voltage is determined by the circuit parameters. The still-moving or already fully-open breaker contacts must be able to withstand the recovery voltage. The most severe stress for the open contact gap is the *initial peak* and the *rate of rise* (kV/μs) of the recovery voltage.

If the recovery voltage exceeds the gap insulation, the arc will restrike and current will continue until the next current zero, when interruption will again be attempted. The rate of rise of recovery voltage is a function of the constants of the circuits which supply power through the breaker. The larger the adjacent capacitance to ground before the major inductance limiting the fault current, the slower will be the rise of the recovery voltage. Some breakers modify the recovery voltage characteristics by limiting the current, modifying its power factor, and so on.

10.2.2 Severe Interrupting Conditions

The following severe cases of circuit-breaker switching conditions have to be considered carefully: terminal fault, short-line fault, out-of-phase switching, switching of small inductive currents, switching of capacitive currents, closing on a fault. The chosen examples are typical only; they are uniformly based on similar and simplified network configurations and are restricted to single-phase fault conditions. Other switching conditions may also be important.

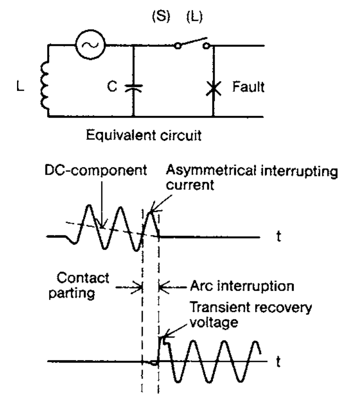


FIGURE 10-64 Principle of terminal-fault interruption, equivalent circuit; typical shape of short-circuit current and transient recovery voltage.

Terminal Fault. After interruption of short-circuit current, the recovery voltage oscillates toward the service frequency driving voltage via an initial peak. The natural frequency is determined by the inductance and capacitance of the driving system (Fig. 10-64).

The dc component of the short-circuit current depends on the time constants of the network components like generators, transformers, cables, and high-voltage lines and their reactances of the zero-sequence and the positive sequence networks. The recovery voltage will accordingly vary depending on the location of the circuit breaker within the network.

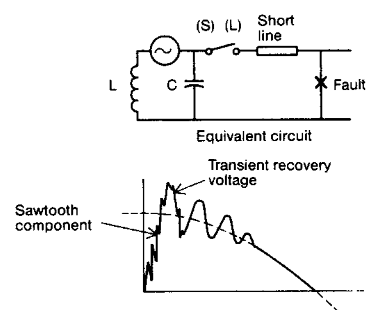


FIGURE 10-65 Principle of short-line fault interruption, equivalent circuit; typical shape of recovery voltage.

Short-Line Fault. In the case of a short-line fault, a section of line lies between the breaker and the fault location (Fig. 10-65). After the short-circuit current has been interrupted, the oscillation at the line side (*L*) of the breaker assumes a superimposed "saw-tooth" shape. The rate of rise of this line oscillation is directly proportional to the effective surge impedance and the time rate of change of current (di/dt) at current zero. The component on the supply side (*S*) basically exhibits the same waveform as a terminal fault. The circuit breaker is stressed by the difference between these two voltages. Because of the high frequency of the line oscillation, the transient recovery voltage has a very steep initial rate of rise. Since the initial rate of rise increases with increasing rate of current change, the limiting interrupting capability of many breaker designs is determined by the short-line fault.

Out-of-Phase Switching. Two network systems with driving voltages $E1$ and $E2$ are connected via a high-voltage transmission line (Fig. 10-66). Since the circuit is closed via the closed circuit breaker, the resulting driving voltage is equal to the sum of the two system voltages. Driving voltage $E2$ may, for example, exceed voltage $E1$ by the voltage drop across the transmission line. After opening the breaker, the transient recovery voltages of the disconnected networks oscillate independently.

The circuit breaker is stressed by the difference of these two voltages. In the case of disconnection of long lines, the recovery voltage across the breaker could be increased because of the Ferranti effect, where the voltage of the receiving end can be up to 15% higher than the sending end if the line is lightly loaded.

Interruption of Small Inductive Currents. This occurs (see Fig. 10-67) when disconnecting unloaded transformers, reactors, or compensating coils. An arc is produced between the contacts when the circuit breaker is opened. The arc voltage is approximately constant at higher currents, since the arc energy is removed only be convection. With small currents, the arc voltage increases as a result of arc looping and a change in the cooling mechanism.

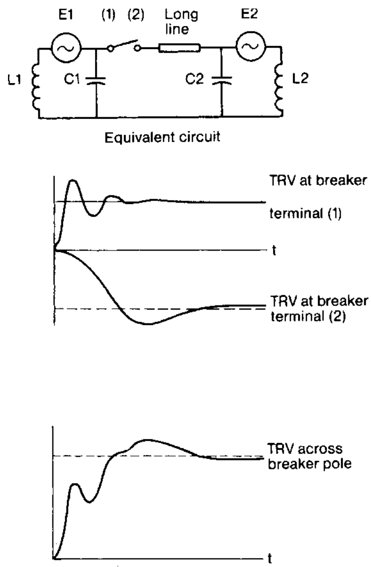


FIGURE 10-66 Principle of out-of-phase switching, equivalent circuit; typical shape of recovery voltage.

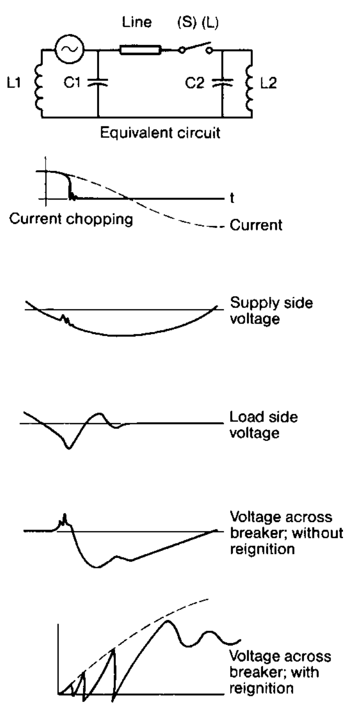


FIGURE 10-67 Principle of small-inductive-current interruption; equivalent circuit; typical shape of current and voltage.

When approaching current zero, the arc current begins to oscillate as a result of interaction with the system; that is it becomes unstable.

As a result of the high oscillation frequency, the current interruption may occur prior to the natural zero passage, can be regarded as instantaneous, and is called *current chopping*. The chopping current is affected not only by the properties of the circuit breaker but also to a great extent by the system parameters.

Energy at the disconnected load side (*L*) oscillates with the natural frequency of the capacitances local to the circuit breaker. The maximum voltage is attained at the moment when all the energy is converted into capacitive energy.

As a result of the resistive losses, the voltage on the disconnected load side decays to zero. During current chopping, the breaker is stressed by the supply-side voltage on one side and by the load voltage on the other side. The supply side voltage is at a maximum, since the load is highly inductive. The load side voltage is the oscillating voltage as the energy exchanges from inductive energy to capacitive energy. This load side voltage will have a high frequency of up to several thousand cycles per second. During this increasing stress, reignition across the breaker may occur.

However, the arc is immediately extinguished again because of the low current and the process begins anew. Hence, the reignition also helps reduce the energy stored in the disconnected circuit.

Interruption of Capacitive Currents. Capacitive currents occur during line drooping as well as during disconnecting unloaded cables or capacitor banks (Fig. 10-68). Although, switching of capacitor banks is regarded as a special application, disconnecting of charged lines is a frequent switching operation.

Current chopping may occur at a low instantaneous current value during interruption of capacitive currents, but this does not lead to overvoltages. After interruption of current, the voltage at the line capacitance (*L*) remains at the peak value of the power frequency voltage, whereas the voltage on the source side (*S*) oscillates about the driving voltage. The difference between the two voltages appears across the circuit breaker with an amplitude of more than double the rated voltage. If the circuit breaker cannot withstand this higher voltage *restriking* may occur. Restriking is similar to closing transmission lines with trapped charge. After restiking, a transient current flows through the circuit breaker, which is of higher frequency than that of the system and which can again be interrupted during the reignition process. After *reextinction*, the line is charged to the potential of the peak value of the equalizing process, whereas the circuit-breaker terminal on the source side (*S*) recovers to the system voltage. A very high differential voltage appears across the breaker, which may lead to renewed restriking and even switching failures. Restrike-free interruption of capacitive currents is thus of the utmost importance. Basically, the same phenomenon occurs during disconnection of capacitor banks. To determine the voltage stresses of the circuit breaker, however, the grounding condition of the supply system and capacitor bank and the arrangement of the bank have to be taken into account.

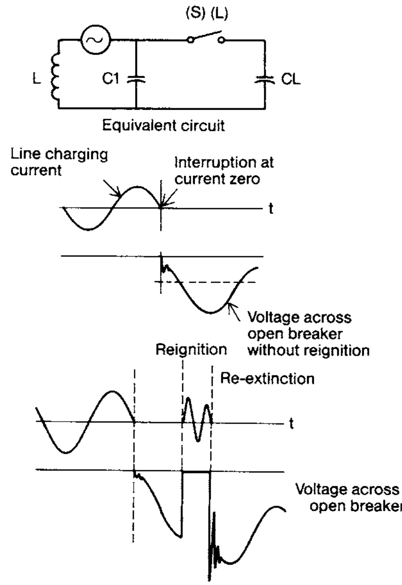


FIGURE 10-68 Principle of capacitive current interruption, equivalent circuit; typical shape of current and voltage.

Closing on a Fault. This (see Fig. 10-69) directs the stress onto the circuit breaker contact system, particularly as regards the electrodynamic and thermal forces. The current and voltage stress is different during closing on (a) symmetrical or (b) asymmetric short-circuit current.

The deciding factor is the moment of contact touch relative to the phase angle of system voltage. In case contact touch and consequently ignition of the arc occurs at the voltage maximum, the short-circuit current will appear symmetrical. The other extreme case takes place with the moment of closing at voltage zero. Here the asymmetrical short-circuit current contains the maximum dc component. A contact system designed for fast closing operation will be subjected to a shorter arcing time and consequently to reduced contact burning when closing on symmetrical currents. Fast operation is therefore not only important for opening but also for circuit-breaker closing.

10.2.3 Ratings and Selection

Voltage Rating and Insulation. Circuit breakers are built for voltage ratings as defined in IEEE C37.04[2] and IEEE C37.06[3]. They have to be dimensioned to withstand the maximum voltages as specified. The rated maximum voltage is the upper limit for operation.

For circuit breakers rated in accordance with ANSI C37.06-1987[4] (or earlier), the range between upper and lower limit is defined by voltage range factor *K*. Current-interrupting capabilities vary within this range in inverse proportion to the operating voltage.

For circuit breakers rated in accordance with ANSI C37.06-1997[4] (or later), the current-interrupting capability is a constant kA value at any voltage equal to or lower than the rated maximum voltage.

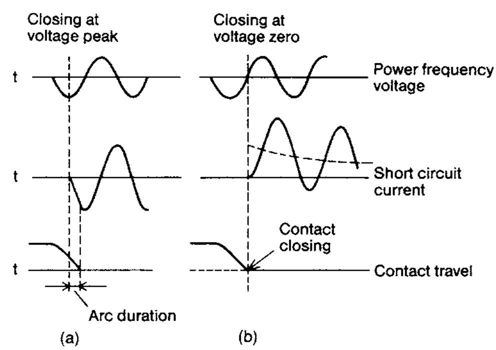


FIGURE 10-69 Stress on contact when closing on a fault, contact travel related to (*a*) symmetric, and (*b*) asymmetrical short-circuit currents.

The insulation level is determined by the rated withstand test voltages specifying the low-frequency voltage (kV, rms) and the impulse voltage (kV, crest). High-voltage breakers must essentially withstand switching surges and both full and chopped-wave lightning impulses. For multiple-break circuit breakers, equal voltage distribution over the series breaks is achieved by grading capacitors paralleled to the interrupting chambers. Coordination between inner and outside insulation, as well as insulation coordination between interrupters and ground insulation, has to be properly designed to prevent flashover inside the breaker or over the open break.

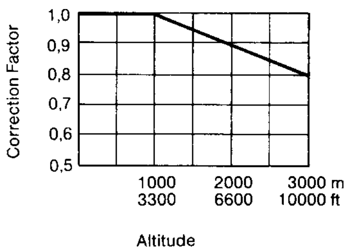


FIGURE 10-70 Altitude-correction factor for voltage ratings.

Outdoor breakers are generally available with special bushings that provide increased creepage distance for installation sites with highly contaminated air. For heavily polluted atmospheres, spray washing of live or deenergized breakers may be an additional measure. Because of the method of design with enclosed ground insulation, the GIS circuit breaker is not influenced by atmospheric pollution.

For installation at altitudes above 3300 ft (1000 m), altitude correction factors have to be applied (Fig. 10-70). The values of rated maximum voltages and insulation levels are multiplied by these factors to obtain the values for the application. The altitude correction factors are as listed in ANSI/IEEE C37.04-1979. Correction factors are under discussion in an IEEE Switchgear committee working group and are expected to change. These factors will be published in IEEE C37.100.1[5].

Particular reference is made to the rating structures and preferred ratings for ac high-voltage circuit breakers per the latest standard revisions of IEEE C37.04[2] and IEEE C37.06[3].

Continuous Current Rating. The *rated continuous current* is the current a circuit breaker has to be able to carry continuously without exceeding a specified temperature rise limit in a specified ambient temperature. Attention must be paid to reduction factors arising from the kind and site of installation, class of insulation material, and electrical endurance requirements.

High-voltage ac breakers must be able to interrupt the rated short-circuit current (kA, symmetrical). They shall be capable of performing the required closing-latching-carrying-interrupting duties in immediate succession. The closing and latching capability (peak kA) is 2.6 times rated short-circuit current (kA, symmetrical) for circuit breakers rated in accordance with ANSI C37.06-1997 (or later). For circuit breakers rated in accordance with ANSI C37.06-1987 (or earlier), the closing and latching capability (rms asymmetrical kA) is 1.6 K times rated short-circuit (kA, symmetrical). The current-carrying capability is determined by the 3-s short time current (kA, rms).

Selection and Application. The proper selection and application of circuit breakers is an extremely important element in the design of an electrical system. Breakers are relied on to separate a defective portion of the system from the remainder to prevent the spread of damage and to permit the good potion to continue in service.

Application conditions and considerations for ac high-voltage circuit breakers are outlined in the latest revision of IEEE C37.010[6], C37.011[7], C37.012[8], and C37.015[9].

Among others, the following criteria have to be considered when selecting a circuit breaker:

System data, such as maximum system voltage, insulation level, short-circuit requirements, and line or cable parameters.

Switching conditions, such as service currents; switching of unloaded transformers, unloaded lines and cables, choke coils, capacitors, generators, and motors; interrupting short-circuit currents and performing special duties like phase opposition, evolving fault, closing on a fault, closing on long lines; duty cycle, reclosing, and operating times.

Service requirements, such as special application for industrial plants, hazardous plants, furnace duty, railway duty, marine duty, maintenance, and operation.

Site of Installation, altitude above 3300 ft, climactic conditions, humidity, wind load, ice, air contamination, space requirements, environmental requirements, earthquake, connection to and function with other switchyard and network components, open installation, or metal-clad switchgear.

10.2.4 Operating Functions

Opening Operation and Duty Cycle. Reaction time and speed of modern breakers has increased to reach standard interrupting times of 2 to 5 cycles, with 2 to 3 cycles being common at high voltage. Interrupting time is measured from energizing of the trip coil until the extinguishing of the arc. The interrupting time during close-open operations may exceed the rated imterrupting time by either $^1/_2$ cycle (for 2 and 3 cycle breakers) or by 1 cycle (for 5 cycle breakers).

The current standard operating duty cycle consists of the following:

$$\text{Open} - T - \text{Close-Open} - 3\ \text{min} - \text{Close-Open}$$

T is defined as either 15 s or 0.3 s depending on whether the circuit breaker is rated for high speed reclosing; this distinction is important in application. Even circuit breakers rated for high speed reclosing must still be allowed a 0.3-s delay to allow for proper recovery of insulation following the initial fault interruption.

For existing oil and air-magnetic circuit breakers, the standard operating duty cycle was: Open − 15s − Close − Open. For additional operations, and/or any close operation in the duty cycle with a time delay of less than 15 s after an opening operation, the interrupting rating and related required capabilities of the oil or air-magnetic breaker have to be derated. All operations within a 15-min period are considered part of the same duty cycle and a duty cycle shall have no more than five opening operations. For guidance on interrupting capability for reclosing service for oil and air-magnetic breakers manufactured after 1960 refer to IEEE C37.010[6]. Circuit breakers manufactured prior to IEEE C37.7-1960[10] have different basis of rating.

Closing Operation. Circuit breakers are designed to perform the closing and reclosing operations as per standard requirements.

When operated to close on long lines, extra-high-voltage circuit breakers require special measures to keep switching overvoltages within specified limits. Such measures may be single or multiple step *closing resistors*, synchronously closing at the moment of voltage zero, or *polarity-controlled-closing*, which means closing during the period of equal polarity at the line and source side of the breaker.

When operated to close on capacitor banks special measures may be taken to limit transient currents and voltages. Such measures may be closing resistors; controlled closing at the moment of voltage zero for grounded wye capacitor banks; or controlled closing on ungrounded wye capacitor banks where the first phase is closed at the moment of voltage zero and the other two phases are closed at a point where the voltage difference between the two phases is zero.

When operated to close on power transformers or shunt reactors special measures may be taken to limit inrush transient currents and transient voltages. Such measures may be single or multiple step closing resistors, or controlled closing at the moment of voltage peak.

The magnitude of overvoltages on energizing and reenergizing is influenced by the nature and variables of the power system. Parameters of supply side and line must be taken into account in order to compute the overvoltages or to determine them using transient network analyzers or transient analysis software, such as electromagnetic transients programs (EMTP), power systems computer aided designs (PSCAD), or alternative transient program. (ATP).

For a summary of the magnitude of overvoltages occurring when energizing high-voltage lines, based on numerous studies and measurements in high-voltage networks, see Table 10-9. Surge arresters may also be used to limit switching overvoltages.

TABLE 10-9 Overvoltages Occurring When Energizing High-Voltage Lines

Prevailing condition	Overvoltage factor (per unit)
1. Line with trapped charge, no compensation, no means of reduction employed	>3
2. Line without trapped charge, no compensation, no closing resistors, or with trapped charge, no closing resistor, but polarity-dependent closing	2.0−2.8
3. Same as no. 2, but with compensation	2.0−2.5
4. Single-stage closing resistors, compensated line	≤ 2.0
5. Two-stage closing resistors, optimum compensation	≤ 1.7
6. Two-stage closing resistors, combined with polarity-dependent closing, or compensation with optimized multistage closing resistors	1.5

Operating Mechanism. Opening and closing of power circuit breakers under service conditions is seldom performed manually, since most breakers are installed in systems designed for remote control providing specific redundancy. Various means of operation are used, such as (1) dc solenoids, (2) solenoids operated from an ac source through a dry-type rectifier, (3) compressed air, (4) high pressure oil, (5) charged spring, and (6) electric motor. *Automatic reclosing* of breakers in overhead line feeders is frequently used to restore service quickly after a line trips out because of lighting or other transitory fault. Instantaneous or time-delay reclosing may be provided with a lockout to prevent more than one to several successive reclosures, as desired. If the fault is cleared before the lockout feature operates, the reclosing device resets itself, permitting a complete cycle of reclosing at a subsequent fault.

The circuit-breaker-operating device has to cope with the increasing requirements in interrupting and current-carrying capability as well as with shorter operating times. Simplicity of design, robustness, and reliability have to ensure safe operation of this vital link between the electrical system controls and the interrupter. The principle of a pneumatic drive is sketched for an extra-high-voltage circuit breaker which functions according to the differential piston principle in Fig. 10-71. A pneumatic interlocking device in connection with the SF_6 gas system ensures that the breaker always remains in the defined open or closed position even on loss of air pressure. Besides opening and closing functions, effective damping of the highly accelerated moving parts is incorporated.

Accessories. Circuit breakers may be equipped with a wide range of accessories, either required, like pressure controls, gas-density monitors, safety valves, and position indicator, or optional, such as a choice of different release, alarms, or auxiliary contacts. To illustrate the importance of accessories for safe and reliable circuit breaker operation, Fig. 10-72 shows the SF_6 gas monitoring system of a high-voltage SF_6 outdoor breaker.

The insulation and breaking capacity of an SF_6 breaker depends on the gas density. It is assumed that the volume remains constant during temperature variations, whereas the pressure of the SF_6 is highly dependant on temperature change. Hence, to monitor the state of the gas, it is logical to supervise not the pressure but the density of the gas.

The density monitor operates according to the principle of a temperature-compensated pressure gage, the characteristics of which correspond to the constant-density line. The SF_6 gas pressure acts on a metal bellows, the movement of which is transmitted by a transfer mechanism with a bimetal disk to the microswitch.

The density monitor is set for the operating pressure. The pressure-temperature diagram shows the standard case for this type of breaker, a minimum pressure of 5 bar, measured at 20°C. The density monitor emits a signal at 5.2 bar, indicating that refilling is necessary. If the pressure drops below 5 bar, operation of the breaker is blocked.

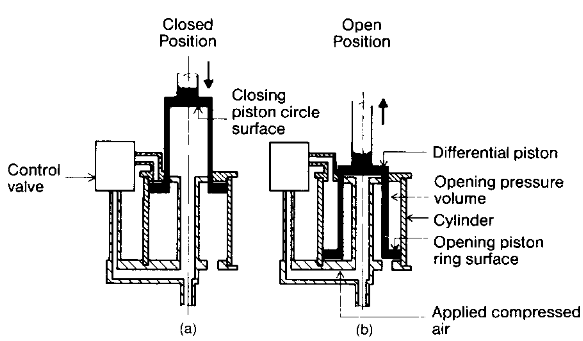


FIGURE 10-71 Principle of the drive system for an SF_6 outdoor breaker: (*a*) closed position; (*b*) open position.

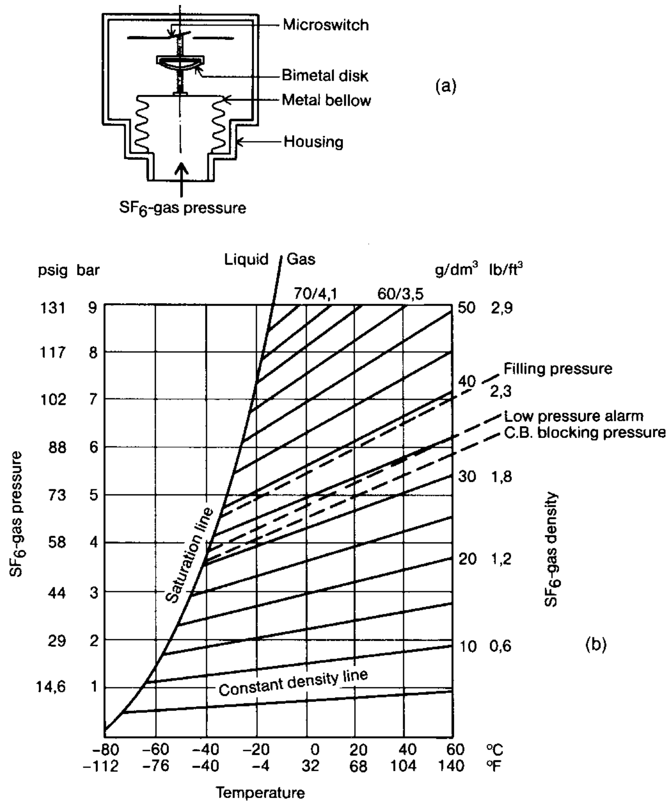


FIGURE 10-72 (*a*) Arrangement and (*b*) pressure-temperature diagram for SF_6 gas-density monitor system for an outdoor breaker. *(Brown Boveri.)*

10.2.5 Testing and Installation

Development and Design Testing of Circuit Breakers. Developing high voltage breakers requires attention to several different technical problems, including withstanding high voltages, short circuit interruption, continuous current capability, mechanical endurance, and environmental robustness. Testing programs are carried out to ensure that the circuit breaker is capable of withstanding these stresses. Test procedures for ac high-voltage circuit breakers are specified in IEEE C37.09[11].

High voltage. The breaker must carry current at high voltage, and be able to withstand transient surges at much higher levels (lightning strikes, for example). Typical test voltage levels are shown in Table 10-10.

Designing for these voltages requires specialized engineering software and knowledge. Dimensions are set for insulating gaps between contacts, and between live parts and the grounded surrounding structures (tanks, shields, etc.). Electromagnetic finite element analysis software is the main tool for this task. SF_6 breakers use pressurized SF_6 gas for insulation and arc quenching. This is because pressurized SF_6 insulates about 15 times better than air, meaning gaps can be much smaller for the same voltage and is vastly better for arc quenching (see Fig. 10-59).

High voltage testing requires "high voltage laboratory" capability with both power frequency "hi-pot" test capability, and also voltage surge ("lightning impulse" and "switching impulse") test capability (see Fig. 10-73).

TABLE 10.10 Typical Test Voltages for Outdoor High-voltage ac Circuit Breakers

Rated maximum voltage kV	Power frequency withstand voltage (1min. dry) kV, rms	Full wave withstand (BIL) kV, peak
15.5	50	110
38	80	200
72.5	160	350
145	310	650
170	365	750
245	425	900
362	555	1300
550	860	1800
800	960	2050

Short-circuit interruption. The most important function of circuit breakers is to interrupt short-circuit currents. This is to protect generators, transmission lines, transformers, and other components of the transmission system. Typical short-circuit requirements of high voltage systems are 25 to 63 kA, though there is an increasing need for 80+ kA. During a short circuit (fault), the circuit breaker is subjected to both high currents and voltages at the same time. Designing for this capability involves engineering simulations and computational fluid dynamic analysis. Testing is difficult, as the momentary test power requirements can exceed even that available on the transmission grid. Therefore, so-called "synthetic" techniques are used for this testing. Synthetic testing involves separate sources for the high current and high voltage, and only combining them during a very brief window of time. This greatly reduces the power requirements of simulating short circuit interruption. A typical high-voltage synthetic power lab is shown in Fig. 10-74.

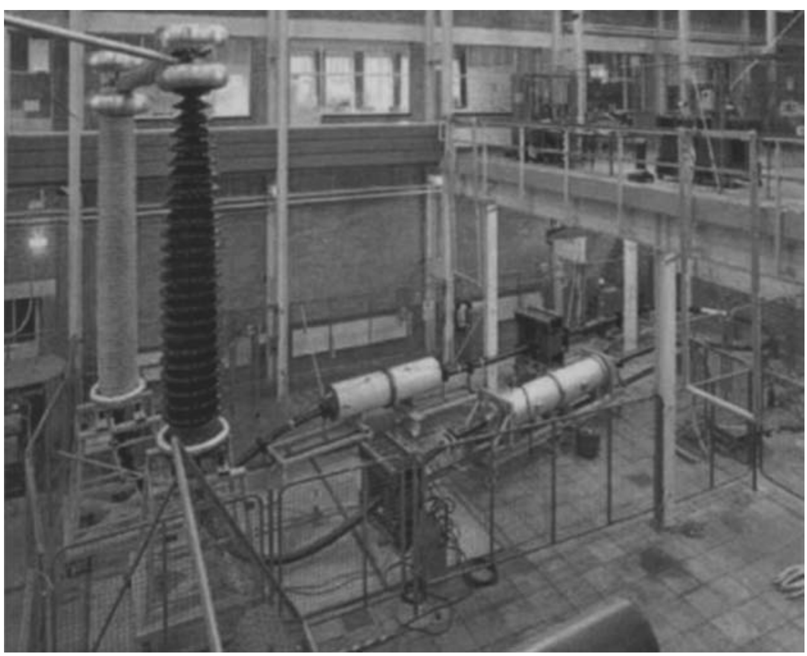


FIGURE 10-73 View of the B.V. KEMA high-voltage laboratory in Arnhem, The Netherlands. (*Courtesy of B.V. KEMA.*)

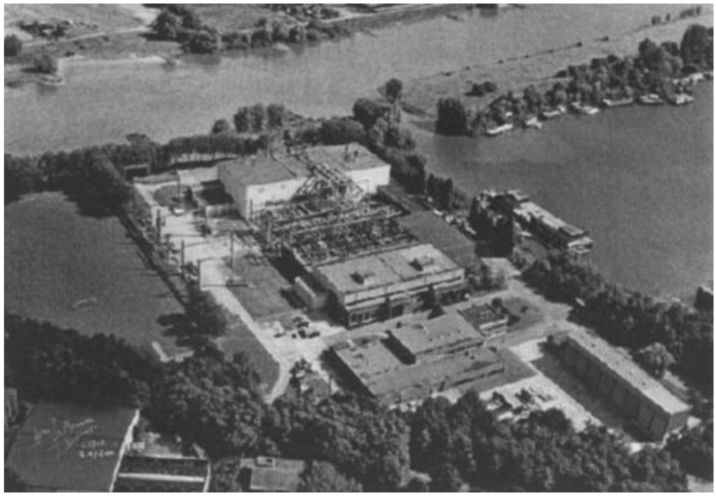


FIGURE 10-74 Aerial view of the B.V. KEMA high-power laboratory in Arnhem, The Netherlands. *(Courtesy of B.V. KEMA)*

The high current is typically supplied from dedicated "short-circuit generators," which are specialized machines in the 1000 to 3000 short-circuit MVA class (see Fig. 10-75). High voltage is supplied from a "high voltage synthetic circuit," which is a combination of capacitor banks, reactors, triggering, circuits, and computer controls (see Fig. 10-76). This synthetic circuit can produce a high voltage waveshape approximating real system recovery voltages.

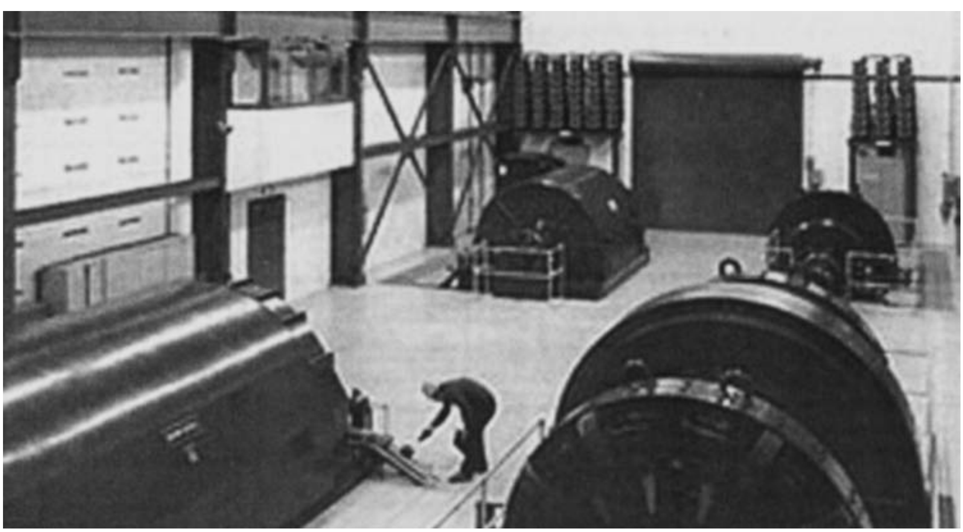


FIGURE 10-75 View of the KEMA-Powertest generator hall with 2,250 MVA and 1000 MVA short-circuit generators. *(Courtesy of KEMA-Powertest, Chalfont, Pa. USA.)*

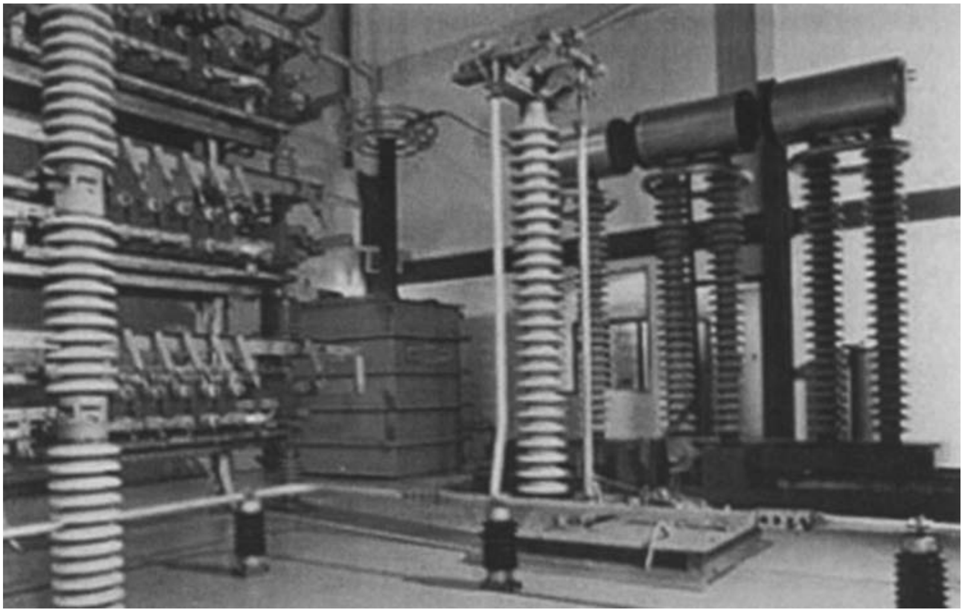


FIGURE 10-76 View of the KEMA-Powertest synthetic test circuit. *(Courtesy of KEMA-Powertest, Chalfont, Pa. USA.)*

In a real world, short-circuit interruption, a transient occurs following interruption which tries to reestablish the arc. This transient voltage must be synthesized by the test circuit and the required wave shapes differ widely between different breaker switching duties. A typical test involves initiating a short circuit with the high current circuit, and then at a "target" current zero simultaneously firing the high voltage synthetic circuit and switching out the high current circuit, (see Fig. 10-77).

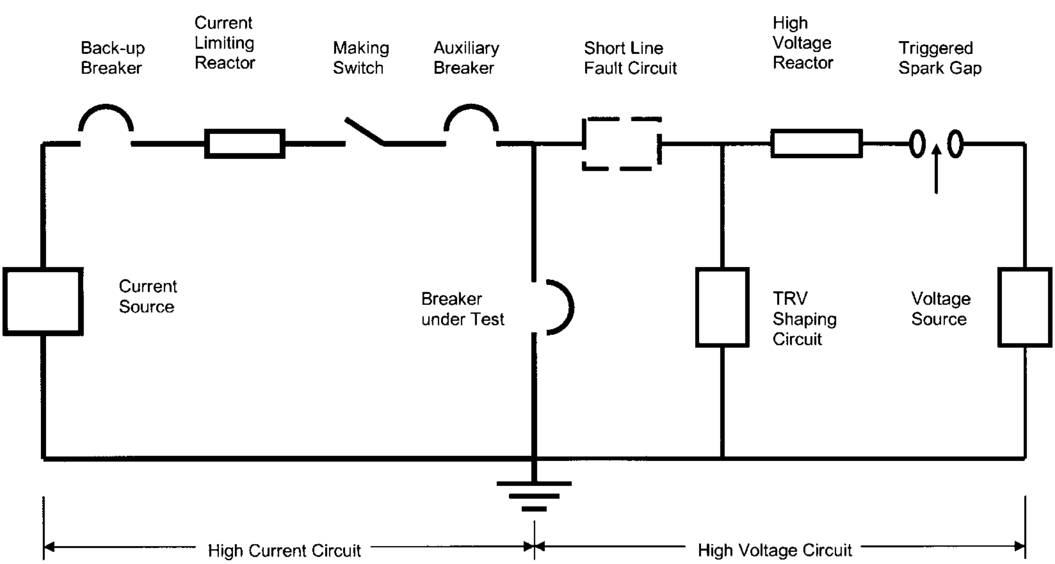


FIGURE 10-77 Parallel current injection test circuit.

The firing of the high voltage synthetic circuit provides correct short-circuit current *di/dt* at the target current zero, and also provides the correct recovery voltage waveshape after successful interruption. Variations of the basic synthetic test circuit are used to simulate different switching and interrupting duties on the power system. Unit testing—testing only one interrupting break of a multi-break circuit breaker—is accepted by all current standards, with some qualification.

Continuous current capability. Circuit breakers must carry a rated continuous load current without exceeding allowable temperature limits. This is demonstrated by testing with a suitable test transformer and voltage/current regulator.

Mechanical endurance. Circuit breakers are intended to operate for many years in service without significant amounts of maintenance. They also must be capable of withstanding many switching operations over that life. Required testing to at least 2000 switching operations is required by standards, though circuit breakers may be required to meet higher values depending on the application and required ratings.

Environmental. Circuit breakers are applied in all temperature zones and in many severe seismic zones. Climatic testing is carried out to verify operation at temperature extremes and to verify the performance of heating systems, when required. For example, SF_6 circuit breakers may require that the insulating SF_6 gas remain heated to a temperature of at least –30°C so that liquification does not reduce the gas density. Reduction of the insulating gas density below allowable levels may reduce insulating and interrupting capability such that the circuit breaker can no longer meet its rated performance (see Fig. 10-72).

Seismic withstand capability is demonstrated either by finite element analysis (FEA) or by shaker table testing, depending on the voltage class. Typically, testing is required at 245 kV and above, with FEA being acceptable below 245 kV. Standard requirements for seismic specification, testing, and application are detailed in IEEE 693[12].

Production Tests. Every circuit breaker is subjected to a series of routine production tests primarily intended to prove design conformance and quality. These tests typically include high voltage power frequency tests ("hi-pot"), mechanical operation and timing tests, fluid/gas leakage tests (when applicable), and control circuit operation verification.

Installation and field tests. These tests are carried out on-site according to specific users or manufacturer's instructions. Modern SF_6 breakers up to 245 kV are usually shipped fully assembled with a slight overpressure of SF_6, thus eliminating evacuation procedures on-site.

Service and Maintenance. With rising system voltages, currents, interrupting ratings, and the requirement for uninterrupted power supply, circuit breaker reliability becomes more and more important. Besides influencing factors of (1) design, (2) quality assurance, and (3) testing, which are mainly a responsibility of the circuit breaker manufacturer, maximum attention must be paid to the maintenance during service. Maintenance instructions for different makes and types of circuit breakers may differ considerably in details and volume, but all strive to obtain maximum breaker reliability despite longer maintenance intervals, smaller inventories of exchange parts, and shorter maintenance hours. Efforts are made to find the easiest way of handling service without influencing neighboring gear and consequently obtaining the lowest service costs. Utility maintenance shafts, standardizing groups, and circuit breaker developers have taken into account these requirements. The various steps from oil breaker to air-blast and finally the SF_6 and vacuum breakers indicate a considerable minimizing of maintenance combined with maximum reliability.

10.2.6 Low-Voltage Circuit Breakers

Application. Air circuit breakers are used on dc and ac circuits for the protection of general lighting, power, and motor circuits.

Distinction is made between various protection classes and different service and ambient conditions. For selection of a breaker, type and rating, operating speed, selectivity with fuses, and high voltage must be taken into account.

Further consideration has to be given to severe or hazardous service conditions like tropical climate or marine- or explosion-proof installations.

Reference is made to IEEE C37.13[13], C37.14[14], and C37.17[15]; UL-489[16]; and ANSI C37.16[17].

Ratings. Standard electrically and manually operated breakers are listed in ratings up to and including 5000 A ac and 12,000 A dc. Electrically operated breakers are available in higher current ratings for special applications. Standard breakers are rated on the basis of a temperature rise on the contacts and terminals not to exceed 50°C above an ambient of 40°C (class 90 insulation). Voltage ratings are 254 to 635 V ac and 250 to 3200 V dc.

The short-time current ratings are based on 3-phase symmetric short-circuit currents; the single-phase short-circuit current ratings are 87% of these values. For details, refer to the latest revisions of ANSI C37.16[17].

Assembly Variations. The breakers are usually installed in a metal-enclosed cubicle for dead-front or drawout type of construction. Metal barriers between breakers and busbars provide increased safety in service.

Hand operation by means of a lever is common, even on large breakers. Electric operation by means of a solenoid or motor mechanisms for 48, 125, or 250 V dc, or 120 or 240 V ac is obtainable on all but the smallest sizes of breakers.

Breakers are supplied with an overcurrent trip mechanism which may be of the instantaneous or the time-delay type, or a combination of both. Trip devices are adjustable over a wide range of ratings. Other trip devices and arrangements may be used, for example, undervoltage trips, shunt trips connected to overvoltage, reverse current, or overcurrent relays.

Multiple-pole circuit breakers are commonly used in practically all capacities, one pole being used for each ungrounded line of a circuit, that is, a 2-pole breaker for a 3-wire grounded circuit or a single-pole breaker for a 2-wire grounded circuit.

Breakers can usually be equipped with auxiliary contacts, alarm contacts, push-button control, position indicator, and key interlock. The widely used drawout type of breaker may be moved into and locked in the connected, test, and disconnected positions and/or completely withdrawn.

Refer to the latest revisions of IEEE C37.13[13], C37.14[14], and C37.17[15], and ANSI C37.16[17].

Air Circuit Breaker. The usual construction of an air circuit breaker (Fig. 10-71) makes use of two fixed terminals mounted one above and the other in a vertical plane, which, when the breaker is closed, are bridged under heavy pressure by a bridging member operated by a system of linkages. Auxiliary and arcing contacts close before and open after the main contacts. The arcing contacts are easily renewable. The breaker is held closed by a latch which may be tripped electrically or mechanically. Modern breakers are trip-free.

Many breakers use a solid bridging member with spring-mounted self-aligning contacts. The contact surfaces are made of silver so that oxidation will not cause excessive resistance and overheating.

Arcing contacts of modern breakers use a silver-tungsten or copper-tungsten alloy which is arc-resistant. The secondary contacts, where used, are usually of copper or silver alloy.

Barriers between poles are generally furnished with breakers on ac and dc circuits 250 V and above, and special arc chutes, quenchers, or deionizing chambers are also used throughout the available lines of air circuit breakers. These devices are made in different forms by different manufacturers and serve to improve the interrupting performance of the breaker and to shorten the arcing time.

Molded-Case Circuit Breaker. This circuit breaker is completely enclosed within a ruggedly constructed molded case of insulating material. It has received wide acceptance in the industry and is particularly adaptable in large buildings and industrial plants. The molded-case circuit breaker, in

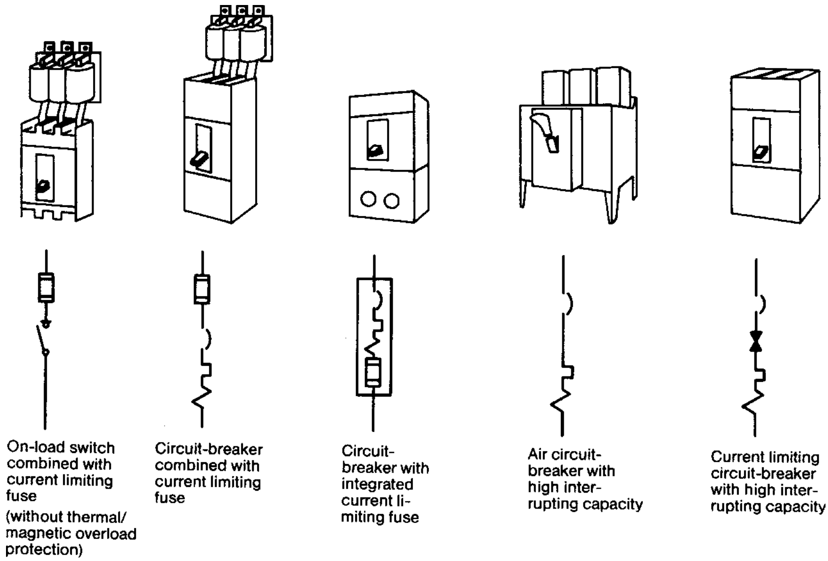


FIGURE 10-78 Methods of current-limiting in low-voltage circuits.

smaller sizes, is adaptable in home lighting circuits where convenience of automatic protection with manual reset of the breaker is desired.

Continuous current ratings range from 15 to 4000 A; the interrupting ratings are from 5 to 45 kA within the standard range. High interrupting ratings up to 200 kA are available.

For details of technical data, application, and accessories, refer to manufacturers' catalogs.

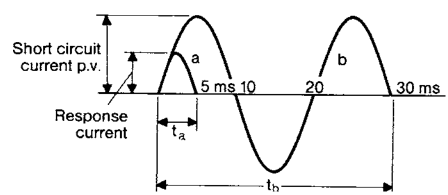


FIGURE 10-79 Current wave (*a*) with limitation, and (*b*) without limitation; t_a = total break time *a*; t_b = total break time *b*.

Current-Limiting Breaker. Low-voltage switchgear is more frequently connected to systems with high or extrahigh short-circuit currents. The standard-range circuit breaker cannot satisfy these requirements. Figure 10-78 outlines different ways to solve the problem. The current-limiting circuit breaker with high interrupting capacity offers a technically sound and economical solution.

Current-limiting breakers operate extremely fast. Interruption takes place within the first half cycle of short-circuit current, so the peak value is not reached. The total break time is less than 5 ms. Figure 10-79 illustrates the current curve, and Fig. 10-80 shows the current-limiting characteristic of a 100-A breaker. With an initial symmetric short-circuit current of 40 kA, the prospective peak value would be 82.5 kA, considering a dc component of 50% and power factor of 0.25. By using a current-limiting breaker, the peak value is limited to about 20 kA. The mechanical stress on the insulation is thus reduced considerably. The contacts in current-limiting circuit breakers are so arranged that the interruption is assisted by the electrodynamic action of the short-circuit current. The higher the short-circuit current, the faster the interruption takes place.

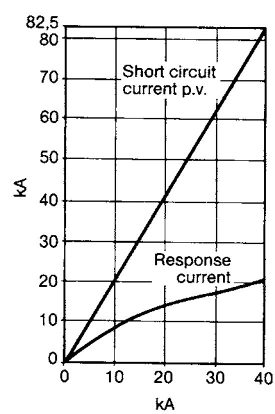


FIGURE 10-80 Current-limiting capability of a motor-protection circuit breaker, 100-A continuous current rating.

Because of the short opening time, the current-limiting circuit breaker, with suitable accessories, can be used to protect power electronic components. Rectifier circuits omitting fuses, for example can be built in this way.

10.2.7 High-Voltage Circuit Breakers

Application. High-voltage circuit breakers have been applied from 1 kV to up to 1100 kV. They are available in either live tank or dead tank design, and may incorporate various insulating and interrupting media as discussed in Sec. 10.2.1.

In today's market, SF_6 circuit breakers dominate above 38 kV, with vacuum circuit breakers being dominant in the "medium" voltage range between 15 and 38 kV. Air magnetic circuit breakers are also common up to 15 kV. Historically, oil circuit breakers were prevalent in these voltage ranges. Even though oil circuit breakers are out of production today, many remain in service.

Indoor breakers in the United States have been historically magnetic air or air-blast types. Today vacuum circuit breakers dominate for indoor application, but there are a few SF_6 types.

For guidance in circuit breaker selection and application, refer to the latest revisions of IEEE C37.12[18], C37.010[6], C37.011[7], and C37.012[8].

Ratings. Standard short circuit currents range from 12.5 to 63 kA, with some applications requiring 80 kA. Commercially available continuous current ratings range from 600 to 5000 A. Preferred ratings are defined in IEEE C37.06[3].

Ratings differentiate between indoor and outdoor service. Preferred ratings are further established for capacitive current switching, dielectric withstand and external insulation, transient recovery voltage capabilities, switching surge factors for line closing, control voltages, reclosing times, and operation endurance capabilities.

Refer to IEEE C37.04[2] and C37.06[3] for the rating structure and preferred ratings.

Oil Circuit Breakers. Oil circuit breakers are out of production today, but many remain in service. These breakers were classified as either dead tank "bulk oil" circuit breakers, or as live tank "minimum oil" circuit breakers. Oil circuit breakers use oil as both an arc quenching and insulating medium, with dead tank "bulk oil" designs using oil as the primary insulation to ground, within a grounded tank.

Dead tank "bulk oil" circuit breakers consist of a steel tank partly filled with oil, through the cover of which are high voltage entrance bushings. Contacts at the bottom of the bushings are bridged by a conducting crosshead carried by a wood or composite lift rod. The breaker typically opens by spring action, separating the interrupting contacts and also further separating an isolation break below the contact system. Accelerating springs are used to increase the speed of opening. In some designs, the crosshead is opened with a rotary motion by springs. Breakers with 3 poles in one tank were made up to 69 kV. Higher voltages had separate tanks for each pole. Figure 10-53 shows a typical dead tank "bulk oil" circuit breaker and Fig. 10-81 shows a typical interrupter.

Minimum oil circuit breakers were developed mainly in Europe to reduce the quantity of oil in circuit breakers. They were manufactured for indoor applications up through 38-kV and outdoor applications up through 800-kV, but were mainly used in the medium voltage range for indoor service. The layout of a medium voltage minimum oil circuit breaker is shown in Fig. 10-82.

Vacuum Circuit Breaker. Progress in high-vacuum technology and breaker development, combined with improved manufacturing and testing methods, has opened a growing area for vacuum breaker application, concentrating, but not limited to voltages up to 38 kV, continuous current ratings up to 3000 A, and covering all standard interrupting ranges. The principal design of a vacuum interrupter is shown in Fig. 10-83.

Two contacts are mounted on an insulating envelope from which virtually all air has been evacuated. One contact is stationary, the other movable. Vacuum interruption has the inherent advantage of moving a lightweight contact only a very small distance in an almost perfect dielectric medium. This results in safe, quiet, and fast switching or interruption of load or fault currents.

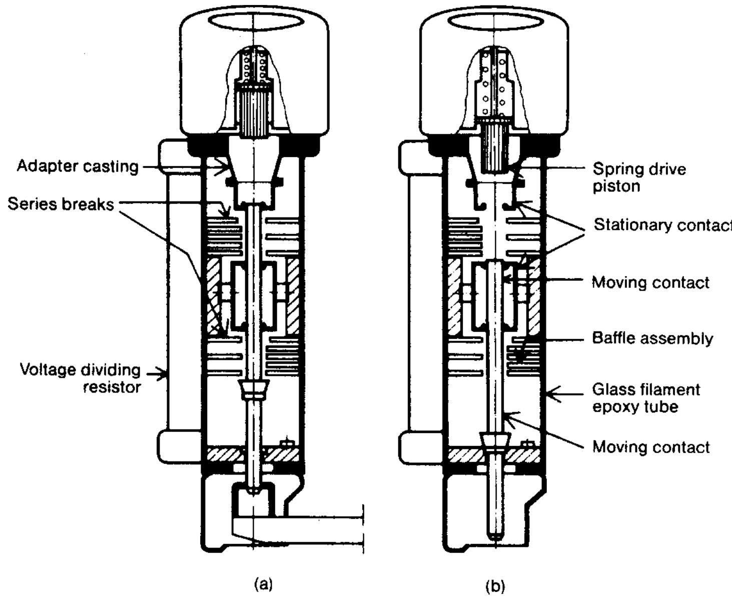


FIGURE 10-81 Details of a 161-kV outdoor oil circuit breaker interrupter: (*a*) closed position; (*b*) open position.

The moving contact is opened up to full gap distance by means of a driving mechanism. A metal-vapor arc discharge thus occurs in the contact gap through which the current flows until the next current zero. The arc is quenched at current zero.

The metal-vapor plasma is fully deionized within a few microseconds by diffusion and recombination so that the conduction path very quickly recovers its dielectric strength. Figure 10-84 shows details of a horizontal-drawout vacuum circuit breaker. One or more interrupters may be utilized in

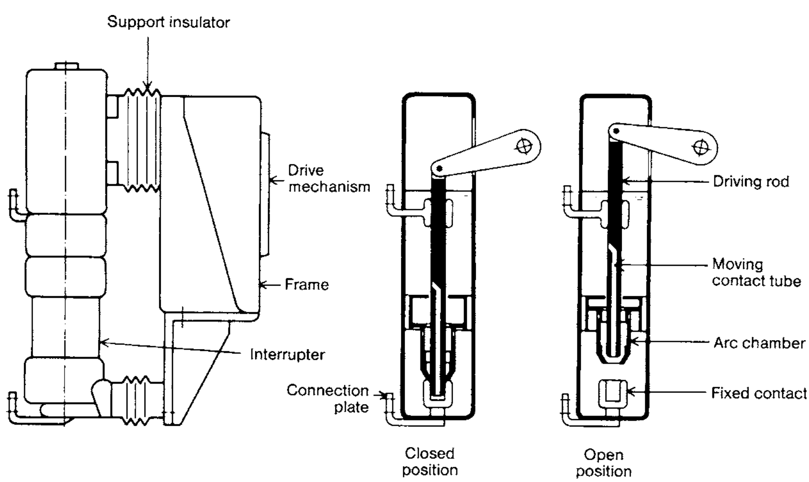


FIGURE 10-82 Outline and interrupter details of a 15-kV, 3-pole minimum-oil circuit breaker. *(Courtesy of ABB T&D Company, Inc.)*

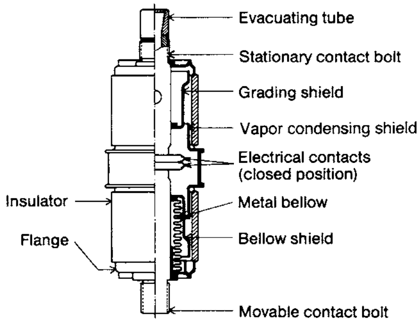


FIGURE 10-83 Partial section of vacuum interrupter 23 kV, 2000 A, 21 kA. *(ABB T&D Company, Inc.)*

series per pole. Vacuum interrupters may additionally be protected against outside influences by an insulating casing. They may also be fitted with hand- or motor-charged stored-energy-operated mechanisms.

Because of their fast closing and opening times, vacuum circuit breakers are particularly suitable for autoreclosure and synchronizing duty. Breaking of the short-circuit currents with very steep initial rise of transient recovery voltage is possible due to restoration of the dielectric strength of the contact gap within a few microseconds. The steep rise of dielectric strength over the whole current range offers a high capacitive-current-switching capability. Switching of unloaded transmission lines and cables can therefore reliably be performed.

Magnetic Air Circuit Breaker. Magnetic air circuit breakers are no longer manufactured. This type of circuit breaker is usually stored-energy mechanism-operated and interrupts the main circuit in the normal atmosphere under the influence of a strong magnetic field which acts to force the arc deep into a specially designed arc chute (see Fig. 10-85). Solenoid operating mechanisms are available. The arc chute cools and lengthens the arc to a point where the arc cannot be maintained by the voltage of the system, and interruption is accomplished. The zone between the main contacts is clear of ionized air by the time interruption is obtained in the arc chute, and so restiking at this point is no problem. Since the magnetic effect is not great at low currents such as small load, transformer magnetizing, and cable-charging current, all designs use an air-pump "puffer" actuated by the operating mechanism that blows a blast of air across the arc and thereby ensures its entering the arc chute and giving rapid interruption at the low-current values also. When the circuit breaker is opened, the arc transfers from the main arcing contacts to fixed arc runners which are within the arc chute. The magnetic field is produced by coils in the main-current circuit, in some cases wound around a magnetic core which magnetizes soft-iron plates in the sides of each arc chute. Some designs do not require an iron core.

Magnetic air breakers were available in any of the ratings of Table 2 of ANSI Standard C37.06-1987 (or earlier) through 15- kV. All were designed for use in metal-clad enclosures. Figure 10-86 shows the horizontal-drawout type of breaker in metal-clad enclosure. Although the design shown is for indoor use only, the same circuit breakers are placed in weatherproof housings for outdoor

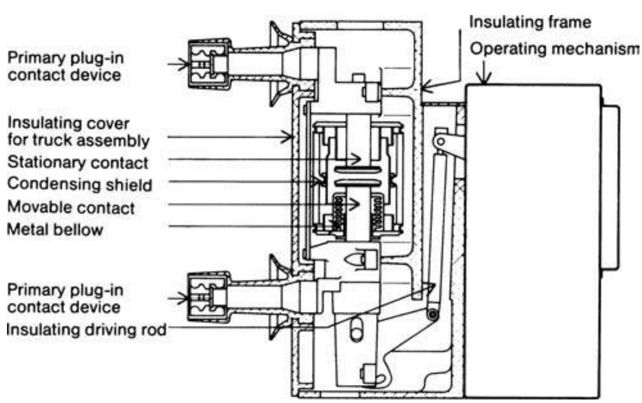


FIGURE 10-84 Outline and interrupter details of a 15-kV horizontal-drawout vacuum circuit breaker 2000A, 28 kA. *(ABB T&D Company, Inc.)*

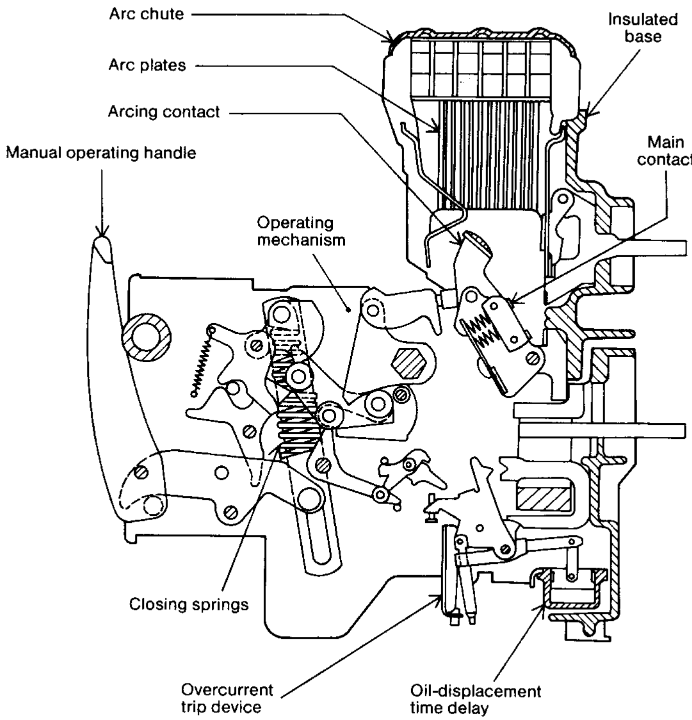


FIGURE 10-85 Typical low-voltage air circuit breaker with magnetic air chutes; breaker in the open position.

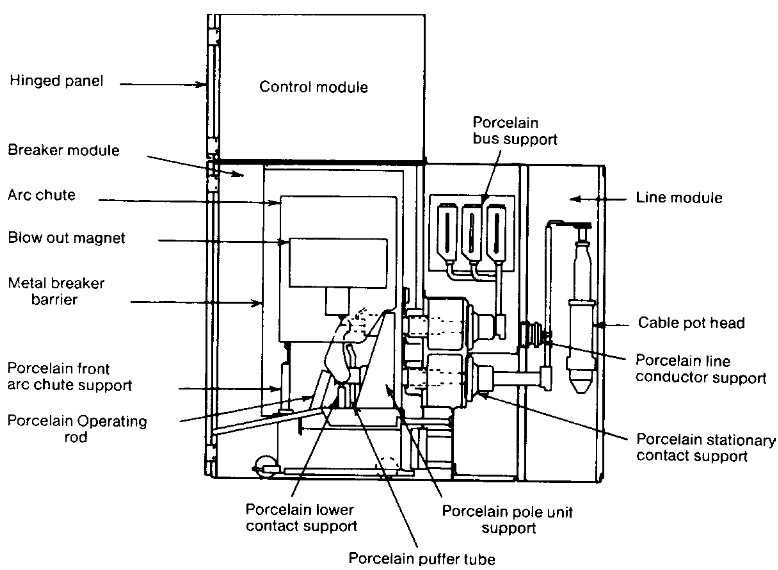


FIGURE 10-86 Horizontal-drawout, metal-clad magnetic circuit breaker in service position.

service. When they are so used, suitable heaters are put in the housings to avoid internal moisture condensation.

Air-Blast Circuit Breakers. Air-Blast circuit breakers are no longer manufactured. These breakers fulfilled the heavy-duty requirements of circuit breakers in high-voltage systems. They have been used to provide the indoor ratings up to 38 kV. They were, however, mainly used in outdoor applications up to 800 kV. Today they have been replaced by SF_6 technology at most ratings.

Air-blast circuit breakers have been used for special applications as (1) generator breakers with continuous current ratings of up to 42 kA and higher, (2) furnace breakers with an extra high number of switching operations (20 to 50 duty cycles per day), and (3) extra high interrupting currents. Air-blast circuit breakers were usually fixed-mounted, but a variety of breaker types were available truck-mounted for application in drawout metal-clad switchgear. All air-blast circuit breakers make use of dry and clean air compressed to a certain pressure, which may differ for the various make and types of breakers. The compressed air is used to operate the breaker as well as to serve as the medium for arc quenching and insulation.

Continuous current ratings up to 5000 A were possible. Total breaking time of 2 cycles (from energizing of trip coil until are extinguishing) was standard; special designs may allow for even shorter breaking time. Some 69-kV breakers are equipped with sequential isolators, but the bulk of designs did not integrate the isolator to form part of the circuit breaker. Some older designs employed separate chambers for opening and closing operation, but later air-blast breakers perform opening and closing with the same contact system. Closing resistors and/or, with some designs, opening resistors were often used. Equal voltage distribution over the multiple breakers of one pole was usually achieved by parallel capacitors.

Generator Circuit Breakers. Generator circuit breakers represent another class rated for very high continuous currents and short-circuit currents, typically at generator voltages. Generator breakers are incorporated into generator bus ducts and can include other switchgear components for measuring current, detecting faults, and grounding.

Generator breakers are available up to 50 kA nominal current and up to 220 kA interrupting current. Two technologies are employed—air blast at the higher ratings (see Fig. 10-87) and SF_6 self-blast at the lower and medium power levels (up to 120 kA). For nominal currents above 20 kA, the generator breaker is usually equipped with a forced cooling system, using water, for example. Generator breakers have been available since the 1960s.

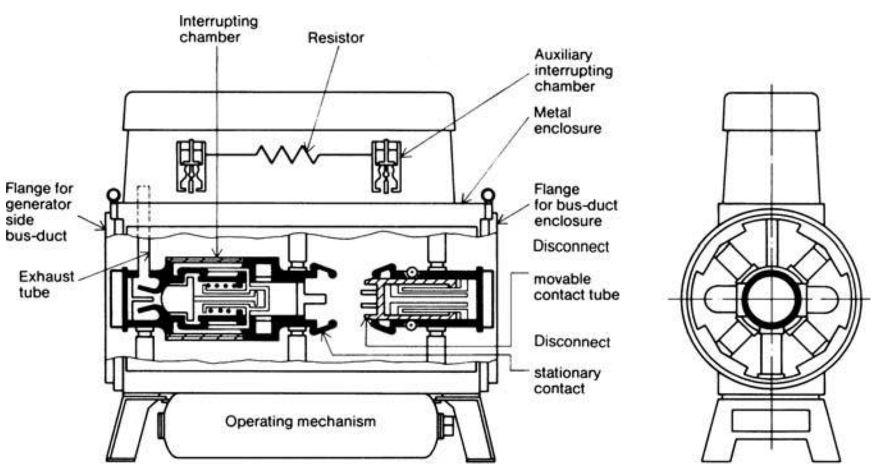


FIGURE 10-87 Outline and interrupter details of a generator air-blast circuit breaker-type DR, 36 kV, up to 50 kA with forced cooling, 200 kA.

Advantages of using generator breakers include the following:

Reduced station cost by eliminating station transformers and increasing station layout flexibility.

Simplification of operation, especially during commissioning and recommissioning; this is because the generator can be handled as a separate unit, isolated from the main and unit transformers.

Fault protection between the generator and transformer. Two zones of protection are created and generator faults are cleared by the opening of the generator breaker alone.

Unbalanced load protection of the generator.

Protection of the generator from transformer faults.

Reliability/availability increase.

Historically, generator circuit breakers have been of air-blast design with pneumatic operators. This is the technology still used today for large nuclear and fossil fuel power plants (up to 1500 MW), and large pumped storage installations. The design has a tubular housing and is horizontal.

Newer designs utilize SF_6 self-blast technology and hydraulic operators. These are rated for application to smaller power plants (gas turbine/cogen, for example) from 60 to 400 MW and smaller pumped storage installations.

SF_6 Circuit Breakers. SF_6 gas has proven to be an excellent arc quenching and insulating medium for circuit breakers. SF_6 is a very stable compound, inert up to about 500°C, non-flammable, non-toxic, odorless, and colorless. At a temperature of about 2000K SF_6 has a very high specific heat, and high thermal conductivity, which promotes cooling of the arc plasma just before and at current zero, and thus facilitates quenching of the arc. The electronegativity behavior of the SF_6, that is, the property of capturing free electrons and forming negative ions, results in high dielectric strength and also promotes rapid dielectric recovery of the arc channel after arc quenching. SF_6 breakers are available for all voltages up to 1100 kV, continuous currents up to 5000 A for conventional breakers (higher for generator breakers), and short-circuit interruption up to 80 kA.

SF_6 breakers of the indoor type have been incorporated into metal-clad switchgear (see Fig. 10-88). Outdoor designs include both dead tank (see Fig. 10-54) and live tank circuit breakers (see Figs. 10-56 and 10-57).

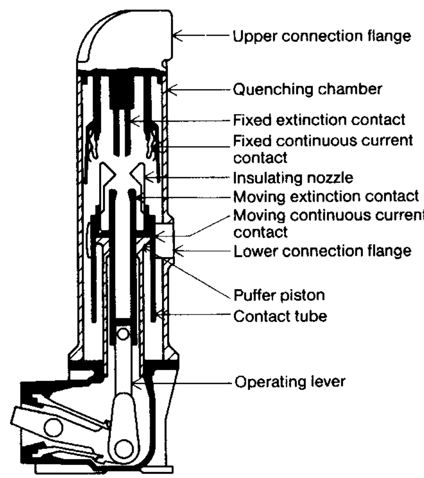


FIGURE 10-88 Section of a SF_6 puffer-piston indoor circuit breaker, 23 kV.

Over the years, SF_6 circuit breakers have reached a high degree of reliability; thus they can cope with all known switching phenomena. Their closed-gas system eliminates external exhaust during switching operations and thus perfectly adapts to environmental requirements. Their compact design considerably reduces space requirements and building and installation costs. In addition, SF_6 circuit breakers require very little maintenance.

All ratings are economically satisfied by the modular design. Each pole is equipped with one or more interrupters; stored energy, spring, hydraulic, or pneumatic driving mechanisms are provided for each pole or 3-pole unit. Gas-density monitors are standard.

Figure 10-89 illustrates the opening sequence of a typical puffer breaker. In the closed position, the current flows over the continuous current contacts and the complete volume of the breaker pole is under the same pressure of SF_6 gas.

The precompression of the SF_6 gas commences with the opening operation. The continuous current contacts separate and the current is transferred to the arcing contacts.

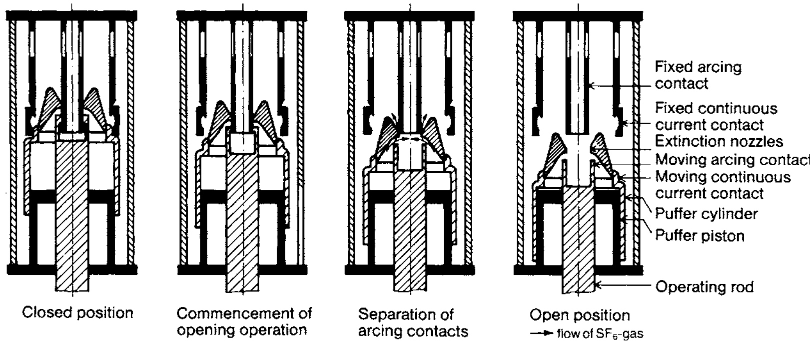


FIGURE 10-89 Principle of SF_6 puffer-type interrupter showing four positions during opening operation.

At the instant of separation of the arcing contacts, the pressure required to extinguish the arc is reached. The arc produced is drawn and at the same time exposed to the gas, which escapes through the ring-shaped space between the extinction nozzle and the moving arcing contact. The escaping gas has the effect of a double blast in both axial directions.

Until the open position is reached, SF_6 gas flows out of the puffer cylinder. The existing overpressure maintains stability of the dielectric strength until the full value of the open contacts at the rated service pressure is reached.

The *self-blast principle* of interruption is illustrated in Fig. 10-90. In the case of high-current interruption, arc energy heats the gas, resulting in a pressure rise in the static volume (heating volume) V_1. This pressure then quenches the arc at an ensuing current zero. In the low-current case an auxiliary puffer (volume, V_2) generates sufficient pressure for interruption. Necessary force requirements for the mechanical system are therefore drastically reduced.

All ancillary equipments, including the oil pump and accumulator associated with the drive, form a modular assembly that is mounted directly on the circuit breaker, thus eliminating installation of piping on the site. The *metal-enclosed GIS breaker* is provided with the necessary items to fit into the substation arrangement (see Fig. 10-91). The main equipment flanges of the breaker are fitted with contact assemblies to accept the isolator moving contacts. Other equipment modules can be coupled to the same flanges. On the fixed-contact end of the circuit breaker, provision is made for coupling two modules, facilitating the mounting of an extension module to connect the second busbar isolator.

Dead tank SF_6 breakers typically employ gas-filled *bushings*, illustrated in Fig. 10-55. Such bushings are usually integral to the circuit breaker itself and are not interchangeable with other apparatus bushings. Electrical grading is provided by a lower throat shield. Ring-type bushing current transformers are located at the base of the bushing. Potential taps are not generally available in SF_6 bushings because of the lack of a capacitive grading structure.

Porcelain alternatives, such as composites, have been used to provide greater safety (explosion resistance), easier handling (lighter and nonbrittle), seismic performance (lighter and stronger), and pollution performance.

Current transformers (CTs) for dead-tank breakers are of the ring-type bushing design. Outdoor breakers of the live-tank layout are generally provided with freestanding CTs of the paper-oil-insulated or SF_6 design (see Fig. 10-57).

For oil-filled CTs, hermetical seal of oil is either of the fixed design with gas cushion or of the pressure-free bellow type. Up to six magnetic cores can be provided per CT unit, generally in multiratios for 5- or 1-A secondary by means of secondary taps. Primary-current ratings up to 2000 A normally employ the wound-type design with two or more turns. Higher primary currents up to 6000 A require the inverted or head design, with a straight tube as single-turn primary winding and the core

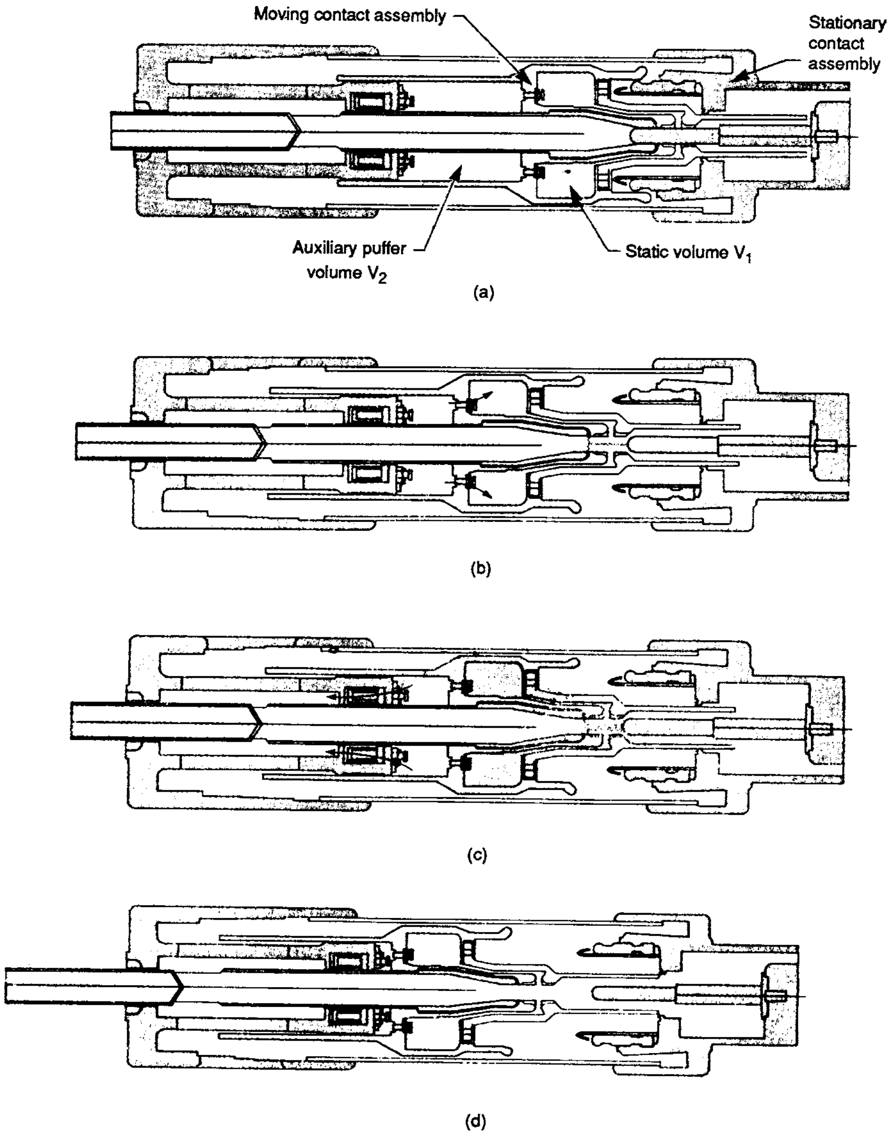


FIGURE 10-90 Self-blast principle of interruption: (*a*) full-closed position; (*b*) low-current interruption; (*c*) high-current interruption; (*d*) full-open position.

and secondary-winding assembly arranged at the CT top to limit temperature rise and to increase the mechanical withstand capability of the CT. The latter design has its full main insulation on the secondary winding.

Freestanding CTs are available for all output and accuracy requirements for modern system relaying and measuring for voltages up to 800 kV. For the upper voltage ranges, freestanding CTs are normally provided with separate potential layers. The CTs are generally dimensioned for the same dielectric and mechanical characteristics chosen for the related circuit breaker.

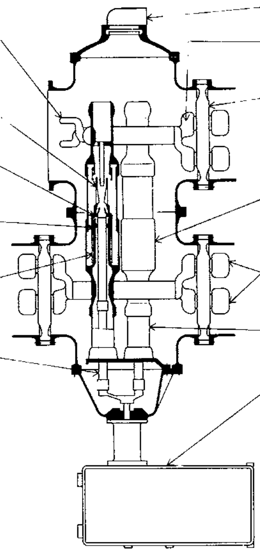


FIGURE 10-91 Section of a 145-kV SF_6 circuit breaker for gas-insulated substation (GIS) type ELK.

REFERENCES

1. IEEE C37.100, *Standard Definitions for Power Switchgear.*
2. IEEE C37.04, *Standard Rating Structure for AC High-Voltage Circuit Breakers.*
3. IEEE C37.06, *Standard for AC High-Voltage Circuit Breakers Rated on a Symmetrical Basis-Preferred Ratings and Related Required Capabilities.*
4. ANSI C37.06, *Standard for AC High-Voltage Circuit Breakers Rated on a Symmetrical Basis-Preferred Ratings and Related Required Capabilities* (Versions of C37.06 prior to 2003 were done by NEMA working groups. In 2003, NEMA transferred ownership of C37.06 to IEEE).
5. IEEE C37.100.1, (This document was being balloted at the time this chapter was written).
6. IEEE C37.010, *Application Guide for AC High-Voltage Circuit Breakers.*
7. IEEE C37.011, *Application Guide for Transient Recovery Voltage for AC High-Voltage Circuit Breakers.*
8. IEEE C37.012, *Application Guide for Capacitance Current Switching for AC High-Voltage Circuit Breakers.*
9. IEEE C37.015, *Application Guide for Shunt Reactor Switching.*
10. IEEE C37.7-1960,
11. IEEE C37.09, *Standard Test Procedure for AC High-Voltage Circuit Breakers.*
12. IEEE 693, *Recommended Practice for Seismic Design of Substations.*
13. IEEE C37.13, *Standard for Low-Voltage AC Power Circuit Breakers Used in Enclosures.*
14. IEEE C37.14, *Standard for Low-Voltage DC Power Circuit Breakers Used in Enclosures.*
15. IEEE C37.17, *Standard for Trip Devices for AC and General Purpose DC Low-Voltage Power Circuit Breakers.*
16. UL-489, *Standard for Safety Molded-Case Circuit Breakers, Molded-Case Switches, and Circuit-Breaker Enclosures.*
17. ANSI C37.16, *Standard for Low-Voltage Power Breakers and AC Power Circuit Protectors-Preferred Ratings, Related Requirements and Application Recommendations.*
18. IEEE C37.12, *Guide for Specification of AC High-Voltage Circuit Breakers.*

BIBLIOGRAPHY

Bibliography of switchgear literature: IEEE committee report, IEEE Transactions on Power Delivery, vol. 5, Issue 1, Jan. 1990, Page(s):177–188.

Bibliography of switchgear literature: IEEE committee report, Veverka, E.F.; Schmunk, E.W.; McCall, L.V.; IEEE Transactions on Power Delivery, vol. 10, Issue 2, April 1995, Page(s):824–844.

Bibliography of switchgear literature. IEEE Committee report, Glinkowski, M.T.; Schmidt, L.; Veverka, E.F.; IEEE Transactions on Power Delivery, vol. 13, Issue 1, Jan. 1998, Page(s):135–156.

Bibliography of switchgear literature, 1992–1996, IEEE/PES Switchgear Committee web page, http://www.ewh.ieee.org/soc/pes/switchgear/index.htm.

Bergman, W.J.B.; 2001 IEEE/PES. Transmission and Distribution Conference and Exposition, vol. 2, 28 Oct.-2 Nov. 2001, Page(s):1071–1076.

Boggs, S. A., Chu, F. Y., and Fujimoto, N., eds., *Gas-Insulated Substations—Technology and Practice*, Pergamon Press, 1986.

Browne, T. E. Jr., ed., *Circuit Interruption—Theory and Techniques*, Marcel Dekker Inc., 1984.

Brunke, John H., *Application of Metal Oxide Surge Arresters for the Control of Line Switching Transients,* paper presented at *Insulation Coordination Seminar—CRA Centennial Meeting*, Toronto, Ontario, May 1991.

Bruettner, D. E., Colclaser, R. G., and Wagner, C. L., Thermal Requirements of Resistors Used in Circuit Breakers for Voltage Control, *IEEE Transactions*, vol. PAS-80, 1970.

Canay, I.M., Comparison of generator circuit-breaker stresses in test laboratory and real service condition, IEEE Transactions on Power Delivery, vol. 16, Issue 3, July 2001, Page(s):415–421.

Colcaser, R.G., Berkebile, L.E., and Buettner, D.E., The Effect of Capacitors on the Short-Line Fault Component of Transient Recovery Voltage, *IEEE Transactions*, vol. PAS-90, 1971.

Dufournet, D., Willieme, J.M.; Recent developments in generator circuit breakers, IEEE/PES Transmission and Distribution Conference and Exhibition 2002: Asia Pacific, vol. 1, 6-10 Oct. 2002, Page(s):88–92.

Dufournet, D., Recent evolution of high-voltage SF_6 circuit-breakers, IEE Colloquium on Physics of Power Interruption, 31 Oct 1995, Page(s):3/1–3/3.

Dufournet, D., Montillet, G.F.; Transient recovery voltages requirements for system source fault interrupting by small generator circuit breakers, IEEE Transactions on Power Delivery, vol. 17, Issue 2, April 2002, Page(s):474–478.

Florschem, C. H., *Power Circuit Breaker Theory and Design*, Peter Peregrinus, 1985.

Garzon, R. D., *High Voltage Circuit Breakers: Design and Applications*, Marcel Dekker Inc., 1997.

Glinkowski, M.T., Gutierrez, M.R.; Braun, D.; Voltage escalation and reignition behavior of vacuum generator circuit breakers during load shedding, IEEE Transactions on Power Delivery, vol. 12, Issue 1, Jan. 1997, Page(s):219–226.

Greenwood, Allan, *Electrical Transients in Power Systems*, Wiley-Interscience, 1971.

Hall, W.M., Gregory, G.D.; Short-circuit ratings and application guidelines for molded-case circuit breakers, IEEE Transactions on Industry Applications, vol. 35, Issue 1, Jan.-Feb. 1999, Page(s):135–143.

Hedman, D. E., Johnson, I. B., Titus, C. H., and Wilson, D. O., "Switching of Extra-High-Voltage Circuits II—Surge Reduction with Circuit-Breaker Resistors", *IEEE Transactions*, vol. PAS-83, 1964.

Janssen, A.L.J., Brunke, J.H., Heising, C.R., Lanz, W., CIGRE WG 13.06 studies on the reliability of single pressure SF_6-gas high-voltage circuit-breakers, IEEE Transactions on Power Delivery, vol. 11, Issue 1, Jan. 1996, Page(s):274–282.

Khalifa, M., ed., *High-Voltage Engineering—Theory and Practice*, Marcel Dekker, 1990.

Kimblin, C.W., Long, R.W., Low-voltage power circuit breakers and molded case circuit breakers-a comparison of test requirements, 1999 IEEE Industrial and Commercial Power Systems Technical Conference, 2–6 May 1999, Page(s):7.

Kimblin, C.W., Long, R.W., *Comparing test requirements for low-voltage circuit breakers*, IEEE Industry Applications Magazine, vol. 6, Issue 1, Jan.-Feb. 2000, Page(s):45–52.

Legate, A.C., Brunke, J.H., Ray, J.J., Yasuda, E.J., *Elimination of closing resistors on EHV circuit breakers*, IEEE Transactions on Power Delivery, vol. 3, Issue 1, Jan. 1988, Page(s):223–231.

McCabe, A.K., Seyrling, G., Mandeville, J.D., Willieme, J.M., *Design and testing of a three-break 800 kV* SF_6 *circuit breaker with ZnO varistors for shunt reactor switching*, IEEE Transactions on Power Delivery, vol. 7, Issue 2, April 1992, Page(s):853–861.

Musa, Y.I., Keri, A.J.F., Halladay, J.A., Jagtiani, A.S., Mandeville, J.D., Johnnerfelt, B., Stenstrom, L., Khan, A.H., Freeman, W.B., *Application of 800-kV SF_6 dead tank circuit breaker with transmission line surge arrester to control switching transient overvoltages*, IEEE Transactions on Power Delivery, vol. 17, Issue 4, Oct. 2002, Page(s):957–962.

Nelson, J.H., *Electric utility considerations for circuit breaker monitoring*, 2001 IEEE/PES Transmission and Distribution Conference and Exposition, vol. 2, 28 Oct.-2 Nov. 2001, Page(s):1094–1097.

Peelo, D.F., Polovick, G.S., Sawada, J.H., Diamanti, P., Presta, R., Sarshar, A., Beauchemin, R., *Mitigation of circuit breaker transient recovery voltages* associated with current limiting reactors, IEEE Transactions on Power Delivery, vol. 11, Issue 2, April 1996, Page(s):865–871.

Ribeiro, J. R., and McCallum, M. E., *An application of Metal Oxide Surge Arresters in the Elimination of Need for Closing Resistors in EHV Breakers*, *IEEE Transactions*, vol. PD-4, 1989.

Roybal, D.D., Standards and ratings for the application of molded-case, insulated-case, and power circuit breakers, IEEE Transactions on Industry Applications, vol. 37, Issue 2, March-April 2001, Page(s):442–451.

Ruoss, E.M., Kolarik, P.L., A new IEEE/ANSI standard for generator circuit breakers, IEEE Transactions on Power Delivery, vol. 10, Issue 2, April 1995, Page(s):811–816.

Smith, R.K., Tests show ability of vacuum circuit breaker to interrupt fast transient recovery voltage rates of rise of transformer secondary faults, IEEE Transactions on Power Delivery, vol. 10, Issue 1, Jan. 1995, Page(s):266–273.

Stevenson, William, D. Jr., *Elements of Power System analysis*, McGraw-Hill, 1975.

Steurer, M., Frohlich, K., Holaus, W., Kaltenegger, K., A novel hybrid current-limiting circuit breaker for medium voltage: principle and test results, IEEE Transactions on Power Delivery, vol. 18, Issue 2, April 2003, Page(s):460–467.

Stoving, P.N., Baranowski, J.F. Interruption life of vacuum circuit breakers, ISDEIV. XIXth International Symposium on Discharges and Electrical Insulation in Vacuum, 2000. Proceedings. vol. 2, 18–22 Sept. 2000, Page(s):388–391.

Sweetser, C., Bergman, W.J., Montillet, G., Mannarino, A., O'Donnell, E.J., Long, R.W.; Nelson, J., Gavazza, R., Jackson, R., Strategies for selecting monitoring of circuit breakers, IEEE Transactions on Power Delivery, vol. 17, Issue 3, July 2002, Page(s):742–746.

Swindler, D.L., Schwartz, P., Hamer, P.S., Lambert, S.R., Transient recovery voltage considerations in the application of medium-voltage circuit breakers, IEEE Transactions on Industry Applications, vol. 33, Issue 2, March-April 1997, Page(s):383–388.

Valentine, R.D., A perspective of low-voltage circuit breaker interrupting rating, IEEE Transactions on Industry Applications, vol. 36, Issue 3, May-June 2000, Page(s):916–919.

Wagner, C.L., and Bankoske, J.W., *Evaluation of Surge Suppression Resistors in High-Voltage Circuit Breakers*, *IEEE Transactions*, vol. PAS-86, 1967.

Wagner, C.L., *Circuit Breaker Application*, Westinghouse Printing,1983.

Wilson, D. D., *Series Compensated Lines—Voltages Across Circuit Breakers and Terminals Caused by Switching*, *IEEE Transactions*, vol. PAS-91, Page(s):1050–1056, 1972.

10.3 SWITCHGEAR ASSEMBLIES

By JEFFREY H. NELSON, MICHAEL W. WACTOR and TED W. OLSEN

Definitions of terms used in this section can be found in the IEEE standards and application guides referenced in this section and/or in the *IEEE Authoritative Dictionary of Terms*. The term "low voltage" as used in this section refers to rated voltages up to 1000 V ac and 3200 V dc. The term "medium voltage" as used in this section refers to rated voltages above 1000 V ac up to 38 kV ac.

Switchgear assemblies cover a wide range of low-voltage and high-voltage *structures that are generally factory* assembled and are divided into the following main groups: (1) metal-enclosed low-voltage power circuit breaker switchgear, (2) medium-voltage metal-clad switchgear, (3) metal-enclosed interrupter switchgear, (4) metal-enclosed bus, and (5) switchboards. ANSI/IEEE C37.20.1[1], C37.20.2[2], C37.20.3[3], C37.23[4] and NEMA PB-2[5] apply. Any of these equipment types may be rated as "arc resistant metal-enclosed switchgear" by successfully meeting the requirements of ANSI/IEEE C37.20.7[6].

10.3.1 Metal-Enclosed Low-Voltage Power Circuit Breaker Switchgear

Metal-enclosed low-voltage power circuit breaker switchgear indicates a design which contains low-voltage ac or dc power circuit breakers in individual grounded metal compartments. The circuit breakers can be either stationary or drawout; manually or electrically operated; fused or unfused; and either 3-pole, 2-pole or single-pole, construction. The switchgear may also contain associated control, instruments, metering, protective and regulating equipment as necessary. Definitions, ratings, design and production tests, construction requirements, and guidelines for application, handling, storage, and installation are covered in IEEE C37.20.1[1].

Low-voltage metal-enclosed switchgear is typically installed in industrial plants, utility and cogeneration facilities, and commercial buildings for the protection and distribution of power for loads such as lighting, machinery, motor control centers, elevators, air conditioning, blowers, compressors, fans, pumps, and motors. Low-voltage switchgear is available in ac ratings up to 635 V and 5000 A continuous and in dc ratings up to 3200 V and 12000 A continuous. Short-circuit current ratings are available up to 200 kA.

10.3.2 Metal-Clad Switchgear

The term "metal-clad switchgear" indicates a design providing metal barriers between primary sections of adjacent vertical sections and between major primary sections of each circuit. Primary sections comprise the bus compartment, the primary entrance compartment, the removable element compartment, the voltage transformer(s) compartment, and the control power transformer(s) compartment. Low-voltage control equipment such as metering, relays, instruments, and controls are located in compartments separate from the primary voltage components. To minimize the possibility of communicating faults between primary sections, the barriers between primary sections shall have no intentional openings. Barriers may be provided to segregate the voltage transformers for each polyphase circuit but not to segregate them individually. Where buses penetrate barriers, suitable bushings or other insulation is required. Definitions, ratings, design and production tests, construction requirements, and guidelines for application, handling, storage and installation are covered in IEEE C37.20.2[2].

Circuit breakers are generally the vacuum type, although air-magnetic circuit breakers were used for many years. Circuit breaker disconnection is accomplished by horizontal-drawout design, illustrated in Figs. 10-92 and 10-93, respectively. Interlocks are provided in metal-clad assemblies to prevent disconnecting or connecting the circuit breaker while in the closed position and to prevent breaker operation while moving between disconnected and connected position or vice versa. The metal-clad assembly is equipped with shutters to protect personnel from coming in contact with the high-voltage circuits when the circuit breaker is removed from the cubicle. A circuit breaker test position is standard to allow breaker control with the main contacts (primary disconnecting devices) removed from the primary circuit, but maintaining auxiliary and ground contacts between cubicle and breaker truck.

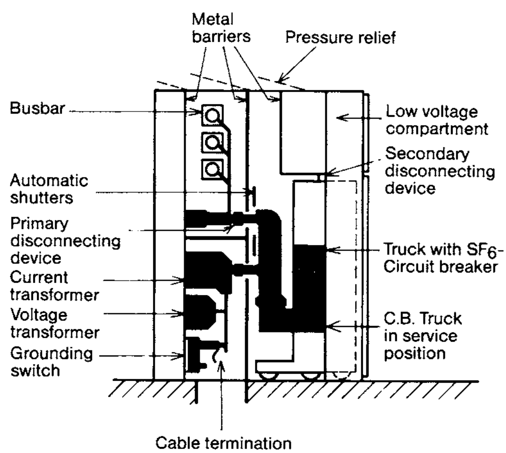


FIGURE 10-92 Side view of a 15-kV metal-clad switchgear unit with horizontal-drawout circuit breaker, interchangeable either minimum oil, SF_6 or vacuum design of equal rating.

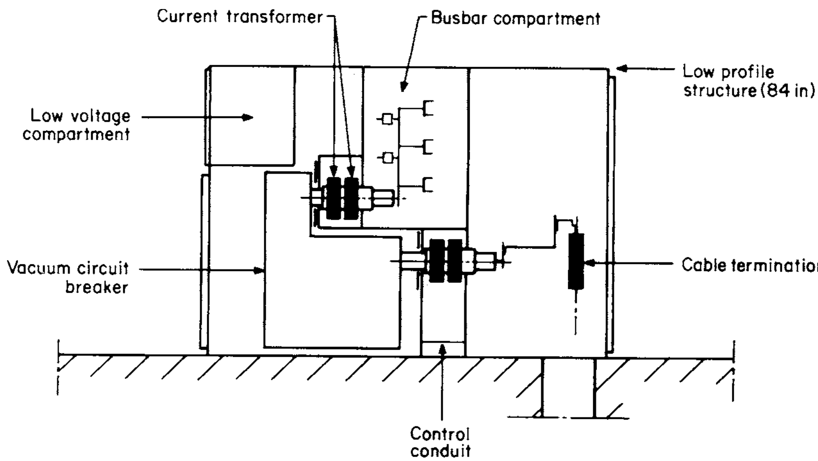


FIGURE 10-93 Side view of a 38-kV metal-clad switchgear unit with horizontal-drawout vacuum circuit breaker, type HKV, up to 3000 A, fan cooled, 22 kA.

Metal-clad switchgear is used for low- and medium-capacity circuits, for indoor and outdoor installations with nominal voltages of 2.4 to 34.5 kV and continuous current ratings typically up to 3000 A. Short-circuit withstand current ratings of the switchgear should equal the ratings of the circuit breaker used.

10.3.3 Metal-Enclosed Interrupter Switchgear

Metal-enclosed interrupter switchgear assemblies include the following equipment as required: interrupter switches, bare bus and connections, selector switches, power fuses (current-limiting or noncurrent-limiting), control and protective equipment, instrumentation, meters, and instrument transformers. The interrupter switches and power fuses may be stationary or removable (drawout). When switches and fuses are removable, mechanical interlocks are provided for proper operating sequence. Also, automatic shutters are provided which cover primary circuit elements when the removable device is in the disconnected, test or removed position.

Definitions, ratings, design and production tests, construction requirements, and guidelines for application, handling, storage, and installation for metal-enclosed interrupter switchgear are covered in IEEE C37.20.3[3].

Metal-enclosed interrupter switchgear is typically used in industrial or institutional environments where continuous load currents are low and frequent switching is not required. Interrupter switches will interrupt load currents up to their rated continuous current capability. Fuses can be installed to provide short-circuit protection. For example, if the interrupter switchgear is connected to other switching equipment fuses can be installed in the connection between the two to prevent an interruption of one assembly for a fault in the other assembly.

Typical applications for interrupter switchgear include main service disconnect, transformer primary and secondary switching, medium voltage switchgear primary and feeder circuit switching. The switching device may be manually operated or motor operated. Motor operated designs are often applied in an automatic transfer scheme.

Metal-enclosed interrupter switchgear is typically available in ac ratings above 1 kV to up to 38 kV and 2000 A continuous current. Short-circuit withstand ratings have to be equivalent to the ratings of the switching and protective equipment used or to the rating of the current transformers used.

10.3.4 Metal-Enclosed Bus

Metal-enclosed bus is an assembly of rigid conductors with associated connections, joints and insulating supports with a grounded metal enclosure. Metal enclosed buses have three basic types of construction: (1) nonsegregated-phase, (2) segregated-phase, and (3) isolated phase.

Rated voltages of ac metal-enclosed bus assemblies range from 635 V through 38 kV, and dc metal-enclosed bus assemblies range from 300 through 3200 V. Definitions, service conditions, ratings, testing, construction requirements, and application guidelines for metal-enclosed bus are covered in IEEE C37.23[4]. An informative guide for calculating losses in isolated-phase bus is also included.

Nonsegregated-Phase Metal-Enclosed Bus. *Nonsegregated-phase metal-enclosed bus* is a type of design in which all phase conductors, with their associated connections, joints, and insulating supports, are enclosed in a common metal housing without barriers between phases, see Fig. 10-94. When associated with metal-clad switchgear, the phase conductors of a non-insulated bus assembly entering the switchgear assembly and connecting to the switchgear bus shall be covered with insulating material equivalent to the switchgear insulation system. Enclosures that are totally enclosed are preferred, but ventilated enclosures can be provided in indoor applications.

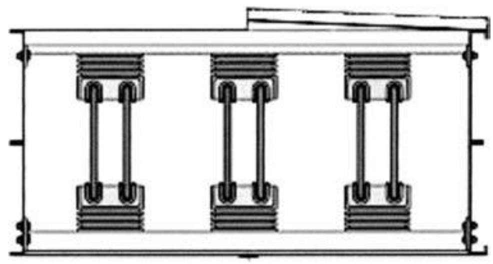


FIGURE 10-94 Typical nonsegregated-phase metal enclosed bus. *(Courtesy of Powell Industries, Inc.)*

Nonsegregated-phase metal-enclosed bus is utilized on circuits which require higher reliability than can be obtained with the application of power cables. Typical applications are the connections between transformers and switchgear assemblies, connections from switchgear assemblies to rotating apparatus, tie connections between switchgear assemblies, connections between motor control centers and large motors, and as main generator leads for small generators.

Preferred continuous self-cooled current ratings for nonsegregated-phase are available up to 12,000 A for 635 V ac and all dc voltage ratings, 6000 A for 4.75 through 15.5 kV, and 3000 A above 15.5 through 38 kV. Short-time withstand current ratings up to 85 kA rms symmetrical are available for ac ratings, and up to 120 kA for dc ratings.

Segregated-Phase Metal-Enclosed Bus. Barriers may be installed between the phase conductors to segregate the conductors and the assembly is then referred to as *"segregated-phase metal-enclosed bus,"* see Fig. 10-95. This design is also used on circuits which require a higher degree of reliability.

Segregated-phase bus is primarily used as generator leads in power plants, but it is also applied in heavy industrial environments and as tie connections in metal-enclosed substations.

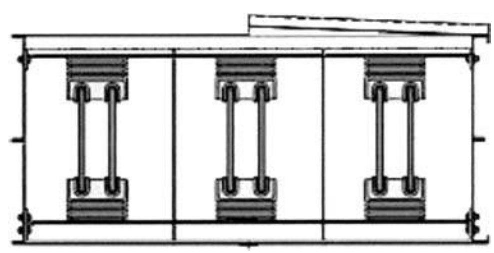


FIGURE 10-95 Typical segregated-phase metal enclosed bus. *(Courtesy of Powell Industries, Inc.)*

Preferred continuous self-cooled current ratings for segregated-phase are available up to 12,000 A for 635 V ac and all dc voltage ratings, 6000 A for 4.75 through 15.5 kV, and 3000 A above 15.5 through 38 kV. Short-time withstand current ratings up to 85 kA rms symmetrical are available for ac ratings, and up to 120 kA for dc ratings.

Isolated-Phase Metal-Enclosed Bus. *"Isolated-phase metal-enclosed bus"* is a type of design in which each phase is enclosed in an individual metal housing, and an air space is provided between the housings. It is considered to be the safest, most practical, and most economical way of preventing phase-to-phase short circuits by means of construction methods. The bus may be self-cooled or force-cooled by circulating air or liquid. Definitions, ratings, design and production tests, construction requirements, and application guidelines for metal-enclosed bus are covered in IEEE C37.23[4].

Briefly, the isolated-phase bus duct has the following features:

1. Proof against contact; locked electrical premises not necessary
2. Faults only in the form of ground faults; protection against fault spreading to more than one phase
3. Field forces, static and dynamic, only between enclosures and conductor, not between phases
4. Protection against contamination and moisture
5. No losses in surrounding conductive material (grilles, railings, concrete reinforcements, lines, etc.)

The range of bus ducts are available up through 38 kV and includes continuous current ratings from about 5 up to 25 kA self-cooled, or 40 kA with forced cooling. The momentary current ratings have to match the rating of attached equipment. With high current ratings, more attention must be paid to the following:

1. Progressive rise of conductor temperature due to skin effects
2. Heating of surrounding conducting material by the magnetic field of conductors
3. High forces on main or component conductors in the event of a short circuit

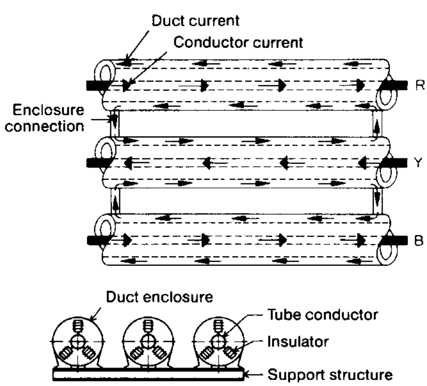


FIGURE 10-96 Three-phase arrangement of an isolated-phase bus duct and principle of enclosure connection; according to Kirchhoff's law sum of conductor currents (+) and sum of duct currents (−) is zero.

In an enclosure with sections of tube insulation (sectional enclosure), eddy currents exist with values as large as the conductor current. These give rise to heat losses, and so the magnetic field of the main conductor is not always compensated for sufficiently. An important technical feature of the bus duct, therefore, is the electrically continuous enclosure. The tubes enclosing each phase have electric conducting joints throughout their length and are short-circuited across the three phases at both ends. The enclosure thus constitutes a secondary circuit to the conductors (Fig. 10-96). The currents in the enclosures reach almost the corresponding conductor currents, depending on the resistance of the duct, but are of the opposite direction. The magnetic field outside the enclosure is almost completely eliminated, and thus there are no external losses or field forces between the phases. Connections to machines and switchgear must be adaptable and removable.

Current transformers for measurement and protection are of the bushing type or are integrated into the bus duct at a suitable place. Voltage transformers can be contained in the bus duct or mounted in separate instrument boards. The same applies to protective capacitors. Care must be taken that branch lines are adequately dimensioned with regard to thermal short-circuit strength.

The reliability of generator bus ducts can be enhanced by employing means to maintain the air pressure in the duct. Although, generally bus ducts are leakproof, the large number of dismantleable joints may cause a slight leakage and might lead to moisture condensation during a plant shutdown. Supplying the bus duct with filtered, precompressed air at slight pressure ensures that the air flow is

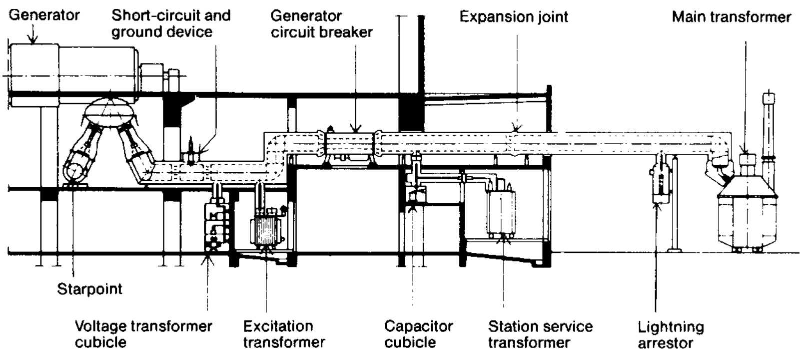


FIGURE 10-97 Generating-plant isolated-phase bus-duct arrangement with generator circuit breaker type DR.

only outward; contamination of the conductors is not possible. Drying the air by precompressing prevents condensation.

Short-circuiting and grounding facilities are required in the bus duct to protect the generator and also for maintenance grounding purposes. Manually positioned links and straps are sufficient for small unit ratings; motor-operated grounding switches are recommended for higher capacities. A typical isolated-phase bus arrangement of a power station including generator circuit breaker is shown in Fig.10-97.

10.3.5 Switchboards

Floor-mounted deadfront "*switchboards*" typically consist of an enclosure, molded case or low-voltage power circuit breakers, fusible or nonfusible switches, instruments, metering equipment, monitoring equipment and/or control equipment, and are fitted with associated interconnections and supporting structures. Switchboards can consist of one or more sections which are electrically and mechanically interconnected. Main disconnect devices can be mounted individually or be an integral part of a panel assembly. Definitions, ratings, design and production tests, construction requirements, and guidelines for application, handling, storage and installation are covered in NEMA PB-2[5].

Switchboards are typically installed in industrial plants, utility and cogeneration facilities, and commercial and residential buildings for the distribution of electricity for light, heat, and power. They are typically available in voltage ratings of 600 V or less, continuous current ratings of 6000 A or less, and short-circuit current ratings up to 200 kA.

10.3.6 Arc-Resistant Metal-Enclosed Switchgear

The term "arc-resistant switchgear" indicates a design in which the equipment has met the requirements of ANSI/IEEE C37.20.7[6]. The switchgear assembly is subjected to an internal arcing fault in key locations throughout the assembly for a specified current level and duration and the equipment performance is evaluated against five basic criteria. The arcing fault is initiated by a small wire placed across the primary conductors which vaporizes when current flows, providing an ionized air path for the arc. The preferred current level for this test is the short-circuit rating of the equipment and the preferred duration for current flow is 0.5 s. The equipment is evaluated for its ability to mitigate conditions which could be hazardous to personnel working

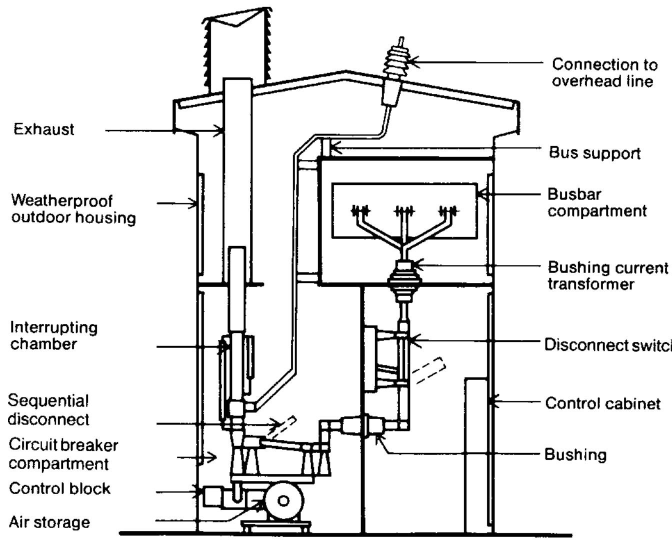


FIGURE 10-98 Metal-enclosed station-type switchgear cubicle for outdoor installation; equipped with a heavy-duty, air-blast circuit breaker, 14.4 kV, 3000 A, 50 kA.

nearby. Definitions, ratings, test requirements, and guidelines for application and installation are covered in IEEE C37.20.7[6].

10.3.7 Station-Type Switchgear

Another type of switchgear assembly that was previously used is "*station-type switchgear*." Station-type switchgear is no longer manufactured, but is briefly discussed here for historical purposes.

The term "station-type switchgear" indicates a design in which the major component parts of a circuit, such as buses, circuit breakers, disconnecting switches, and current and voltage transformers, are in separate metal housings, and the circuit breakers are of the stationary type (Fig. 10-98). Phase segregation in metal-enclosed switchgear is a type of design in which a 3-phase metal housing is divided into three single-phase compartments by means of single metal barriers.

Metal-enclosed station-type switchgear was used in industrial, commercial, and utility installations, generally for voltages of 14.4 to 69 kV, and continuous current ratings up to 5000 A.

REFERENCES

1. IEEE C37.20.1, *Standard for Metal-Enclosed Low-Voltage Power Circuit Breaker Switchgear.*
2. IEEE C37.20.2, *Standard for Metal-Clad Switchgear.*
3. IEEE C37.20.3, *Standards for Metal-Enclosed Interrupter Switchgear.*
4. IEEE C37.23, *Standard for Metal-Enclosed Bus.*
5. NEMA PB-2, *Standard for Deadfront Distribution Switchboards.*
6. IEEE C37.20.7, *Guide for Testing Metal-Enclosed Switchgear Rated up to 38 kV for Internal Arcing Faults.*

BIBLIOGRAPHY

IEEE C37.13, *Standard for AC Low-Voltage Power Circuit Breakers Used in Enclosures.*

IEEE C37.14, Standard for DC Low-Voltage Power Circuit Breakers Used in Enclosures.

IEEE C37.20.4, Standard for Indoor AC Switches (1 kV – 38 kV) for use in Metal-Enclosed Switchgear.

IEEE C37.20.6, *Standard for 4.76 to 38 kV Rated Grounding and Testing Devices Used in Enclosures.*

IEEE C37.21, *Standard for Control Switchboards.*

IEEE C37.22, *Standard Preferred Ratings and Related Required Capabilities for Indoor AC Medium Voltage Switches Used in Metal-Enclosed Switchgear.*

IEEE C37.24, *Guide for Evaluating the Effect of Solar Radiation on Outdoor Metal-Enclosed Switchgear.*

IEEE C37.81, *Guide for Seismic Qualification of Class 1E Metal-Enclosed Power Switchgear Assemblies.*

Bibliography of switchgear literature: IEEE committee report, IEEE Transactions on Power Delivery, vol. 5, Issue 1, Jan. 1990, Page(s):177–188.

Bibliography of switchgear literature: IEEE committee report, Veverka, E.F.; Schmunk, E.W.; McCall, L.V.; IEEE Transactions on Power Delivery, vol. 10, Issue 2, April 1995, Page(s): 824–844.

Bibliography of switchgear literature. IEEE Committee report, Glinkowski, M.T.; Schmidt, L.; Veverka, E.F.; IEEE Transactions on Power Delivery, vol. 13, Issue 1, Jan. 1998, Page(s):135–156.

Bibliography of switchgear literature, 1992–1996, IEEE/PES Switchgear Committee web page, http://www.ewh.ieee.org/soc/pes/switchgear/index.htm.

Bode, K., *Switchgear for 2000 and beyond,* Power Engineering Journal [see also Power Engineer], vol. 14, Issue 6, Dec. 2000, Page(s):286−292.

Bowen, J., Burse, T.A., *Medium-voltage replacement breaker projects,* IEEE Transactions on Industry Applications, vol. 38, Issue 2, March-April 2002, Page(s):584–595.

Bridger, B., Jr., *Comparison of ANSI/IEEE and IEC requirements for metal-clad switchgear,* IEEE Transactions on Industry Applications, vol. 33, Issue 1, Jan.-Feb. 1997, Page(s):216–225.

Bridger, B., Jr., *Application of and rating structure for ground and test devices used in metal-clad switchgear*, Conference Record of the 1988 IEEE Industry Applications Society Annual Meeting, 2–7 Oct. 1988, vol. 2, Page(s):1514–1520.

Bridger, B., Jr., Burse, T.A., Wactor, M.W., *Design considerations for 38 kV metal-clad switchgear using vacuum interrupting technology,* Proceedings of the 1994 IEEE Power Engineering Society Transmission and Distribution Conference, 10–15 April 1994, Page(s):15–20.

Burgin, S.R., Redvers, S.C., Stewart, J.S., *Vacuum technology in partnership with the microprocessor [switchgear]*, Fourth International Conference on Trends in Distribution Switchgear, 1994, 7–9 Nov. 1994, Page(s):18–23.

Conangla, A., White, H. F., *Isolated-Phase Bus Enclosure Loss Factors*, IEEE Transactions on Power Apparatus and Systems, vol. PAS-87, July 1968, page(s): 1622–1628.

Dwight, H. B., *Electrical Coils and Conductors*, New York; McGraw Hill, 1945.

Dwight, H. B., *Some Proximity Effect Formulas for Bus Enclosures*, IEEE Transactions on Power Apparatus and Systems, vol. PAS-83, Dec. 1964, Page(s):1167–1172.

Elgar, E. C., Rehder, R. H., Swerdlow, N., *Measured Losses in Isolated-Phase Bus and Comparison with Calculated Values*, IEEE Transactions on Power Apparatus and Systems, vol. PAS-87, Aug. 1968, Page(s):1724–1730.

Feenan, J., *Low voltage assemblies of switchgear and control-gear comparison of UK and continental European practice,* Third International Conference on Future Trends in Distribution Switchgear, 23–25 Apr. 1990, Page(s):46–50.

Fish, M.W., *When you have to retrofit 15 kV switchgear,* Conference Record of 1994 Annual Pulp and Paper Industry Technical Conference, 20–24 June 1994, Page(s):188–193.

Garzon, R., *The arc terminator,* IEEE Industry Applications Magazine, vol. 9, Issue 3, May-June 2003, Page(s):51–55.

Heberlein, G.E., Jr., Malkowski, C., Jr., Cibulka, M.J., *The effect of altitude on the operation performance of low-voltage switchgear and controlgear components,* IEEE Transactions on Industry Applications, vol. 38, Issue 1, Jan.-Feb. 2002, Page(s):189–194.

Jones, R.A., *Transient recovery voltages for station auxiliary system switchgear,* IEEE Transactions on Power Delivery, vol. 3, Issue 3, July 1988, Page(s):1045–1050.

Kalkstein, E.W., Doughty, R.L., Paullin, A.E., Jackson, J.M., Ryner, J.L., *Safety benefits of arc-resistant metalclad medium-voltage switchgear,* IEEE Transactions on Industry Applications, vol. 31, Issue 6, Nov.-Dec. 1995, Page(s):1402–1411.

Lav, C.T., Staley, D.B., Olsen, T.W., *Practical design considerations for application of GIS MV switchgear,* IEEE Transactions on Industry Applications, vol. 40, Issue 5, Sept.-Oct. 2004, Page(s):1427–1434.

Morton, J.S.; *40 kA 15 kV class vacuum switchgear for power station auxiliaries and distribution systems,* Third International Conference on Future Trends in Distribution Switchgear, 1990, 23–25 Apr. 1990, Page(s):20–25.

Nemoller, A. B., *Isolated-Phase Bus Enclosure Currents*, IEEE Transactions on Power Apparatus and Systems, vol. PAS-87, Aug. 1968, Page(s):1714–1718.

Pihler, J., Ticar, I., Vorsic, J., *Design and development of medium voltage metal-clad switchgear with metal partition walls,* IEEE Transactions on Power Delivery, vol. 18, Issue 2, April 2003, Page(s):475–479.

Swerdlow, N., Buchta, M. A., *Practical Solutions of Inductive Heating Problems Resulting from High-Current Buses*, AIEE Transactions on Power Apparatus and Systems, vol. 78, parat IIIB, Page(s):1736–1746, 1959 (Feb. 1960 section).

Valdes, M.E., Purkayastha, I., Papallo, T., *The single-processor concept for protection and control of circuit breakers in low-voltage switchgear,* IEEE Transactions on Industry Applications, vol. 40, Issue 4, July-Aug. 2004, Page(s):932–940.

Wactor, M., Olsen, T.W., Ball, C.J., Lemmerman, D.J., Puckett, R.J., Zawadzki, J., *Strategies for mitigating the effects of internal arcing faults in medium-voltage metal-enclosed switchgear,* 2001 IEEE/PES Transmission and Distribution Conference and Exposition, vol. 1, 28 Oct.-2 Nov. 2001, Page(s):323–328.

Wilkie, E., *Comparison of ANSI/IEEE and IEC requirements for low-voltage switchgear,* IEEE Transactions on Industry Applications, vol. 40, Issue 6, Nov.-Dec. 2004, Page(s):1656–1664.

Wright, A., *Henry W. Clothier: the `metalclad' man,* Engineering Science and Education Journal, vol. 8, Issue 2, April 1999, Page(s):59–65.

10.4 VOLTAGE REGULATORS

BY CRAIG A. COLOPY

A primary objective of any electrical system is to provide power users with a supply voltage compatible with their utilization equipment. Every electrical device is designed to operate at a certain rated voltage for optimum efficiency and maximum length of service. An ideal electrical supply system provides constant voltage to all users under all loading conditions. No system is ideal; it is economically impractical to attempt an "ideal system" design approach. Today's ideal system is that system providing a voltage supply satisfactory to all utilization equipment, with the most economical use of available regulation equipment.

Several methods of improving the voltage profile on electric transmission or distribution systems are in use. These include transformer load-tap changers, switched and fixed capacitors, and single and three-phase step-voltage regulators. Another new approach, static var control, employing thyristor phase-angle firing control of fixed capacitors is being applied experimentally on several high-voltage transmission systems. Application of single-phase, step-voltage regulators dominates the distribution market to a great extent. They are used in substations having up to 30 MVA loads, as well as being applied to distribution feeders and laterals.

An economical means of voltage regulation is desirable to provide power users with a voltage supply compatible with utilization equipment. The following list describes the effects of unregulated voltage:

1. Resistance loads (electric stoves, irons, water heaters, toasters, etc.)
   a. Low voltage
      (1) Longer heating time.
   b. High voltage
      (1) Shorter life of heating elements.

2. Motor loads (vacuum cleaners, washing machines, fans, refrigerators, etc.)
   a. Low voltage
      (1) Increased slip-increased line current under excited motor. Results: (*a*) decreased efficiency and (*b*) motor overheating.
   b. High voltage
      (1) Overexcited motor-increased torque. Result: possible damage to coupling or appliance.

3. Electronic loads (radio and television)
   a. Low voltage
      (1) Poorer quality of reception on television sets: (*a*) picture not distinct and (*b*) decreased picture size.
   b. High voltage
      (1) Shortens life of electronic components.

4. Illumination loads (incandescent and fluorescent lighting)
   a. Low voltage
      (1) Decreased incandescent lamp efficiency; a 10% decrease in voltage results in 70% normal illumination output.
      (2) If voltage is too low, fluorescent lamps will be inoperative.
   b. High voltage
      (1) Life expectancy of bulbs decreases.

Of course, many benefits of regulated voltage provided to consumers also benefit suppliers by virtue of decreased investment per kVA distributed, increased efficiency of distribution equipment, and increased revenue. The intangible benefit of customer satisfaction must not be overlooked.

10.4.1 Methods of Regulation

The first regulators were induction-type machines. These appeared very early in the development of the electric power industry and were used extensively for a number of years. An induction regulator consists of a rotor and a stator much like a motor. Like step regulators, induction regulators take voltage from the source and add to or subtract from it to hold the load voltage steady. Output voltage is changed by mechanically adjusting the position of the rotor relative to the stator. The rotor does not rotate continuously but has its position changed as required by a small, internal self-contained motor. This motor responds to a signal from a control circuit.

Step-type voltage regulators, the precursors of modern design, were introduced in the early 1930s. The first step-type regulators were developed from the autobooster concept. A 2400-/120- V distribution transformer connected as an autotransformer gives a 5% boost. Adding a center tap to the secondary (series) winding gives two $2^1/_2$% steps. Adding two more taps give four $1^1/_4$ steps. A preventive autotransformer (or bridging reactor) divides these steps in half to give the modern $^5/_8$% steps. In a ±10% regulator, thirty-two $^5/_8$% steps, 16 above and 16 below neutral, provide regulation to all types of loads. Step-type voltage regulators are actually tapped autotransformers. An autotransformer is a transformer in which part of one winding is common to both the primary and the secondary circuits associated with that winding. In other words, the primary (exciter) winding is both electrically and magnetically connected to the secondary (series) winding. The exciter winding is common to both primary and secondary; the series winding is connected in series with load (output) current. Step-type voltage regulators are commonly provided in both single-phase and three-phase styles.

Transformer load-tap changers (commonly referred to as LTC transformers) are actually combination transformers and step-type regulators. The tap-changing mechanism is mounted in an oil-filled compartment which is commonly sealed from the core and coil. Tap changing is accomplished by a gang-operated, three-phase oil switch providing simultaneous regulation of the three-phase voltages.

Fixed and switched capacitors are not voltage regulators in the electrical definition of the device. Capacitors are used to improve the line-voltage profile by connecting a "bank" of capacitors in shunt. These shunt capacitors first improve an otherwise lagging power factor "seen" by a source. The effect of the improved power factor will be reduced line losses and improved regulation. The use of capacitors beyond that required to attain unity power factor will result in a leading component of current in the line inductance, causing voltage on the line to rise. As long as total line current remains lagging, capacitors provide an improved voltage profile. When total current is leading, however, shunt capacitors increase line current (and line losses) and may cause a large voltage rise, resulting in excessively high voltage.

Static var systems (SVS) involve effectively regulating (or fine tuning) the reactive compensation afforded by the shunt capacitors by virtue of phase-angle firing control of thyristors. Basically, the thyristor conduction period will establish the capacitive (leading) current. The control system used is very rapid relative to system fluctuations, permitting optimization of the application. Other features of SVS, especially its use in improving system stability, have made it feasible for exploration on transmission systems. Thus far, SVS application at distribution voltages has been justified only for very large bulk power supplies.

Method of Step-Voltage-Regulator Operation. A typical step-voltage regulator (Fig. 10-99) involves a shunt winding, a series winding, and a bridging reactor or preventive autotransformer.

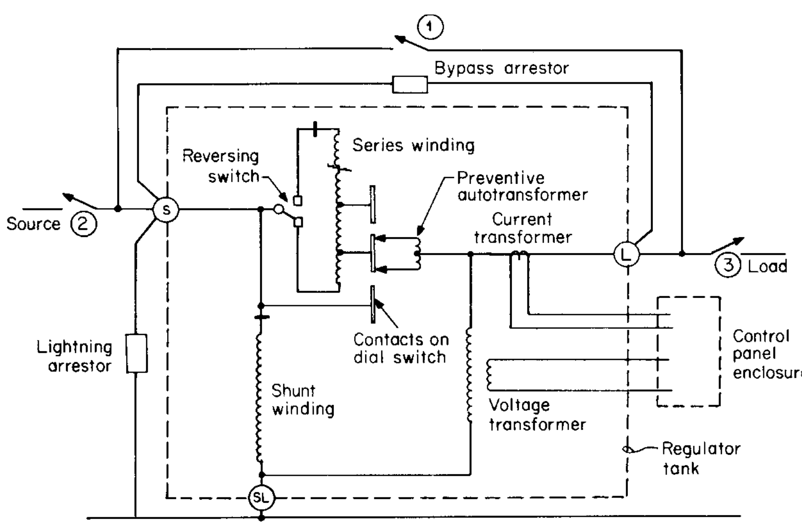


FIGURE 10-99 Wiring diagram of a typical distribution step-voltage regulator showing both external and internal connections, with preventive autotransformer shown on a nonbridging position: (1) bypass switch; (2) source switch; (3) load switch.

Taps, each representing essentially $1^1/_4\%$ voltage, are affixed to the series winding and brought to a specially designed dial switch.

The *main transformer* comprises the *shunt winding* and the *series winding* (Fig. 10-100*a*). The series winding most often is rated at 10% voltage of the shunt winding. Usually eight taps on the series winding are brought to a *dial-switch* assembly as individual contacts, with the voltage difference between contacts being $1^1/_4\%$ voltage.

The terminals of a center-tapped *preventive autotransformer* are able to transit (slide) between the dial switch contacts in a manner which avoids momentary loss of load. As one finger advances to the next step, arcing results because of the inductive nature of the reactor, but load current is maintained (Fig. 10-100*b*) through the finger which does not part contact. When the parting finger remakes on the adjacent contact, a bridging condition is established (Fig. 10-100*c*). A circulating current is established through the preventive autotransformer (PA), or a center tapped bridging reactor, and load voltage is seen to be the average voltage of the taps being bridged. To minimize the arc duration, a *quick-break mechanism* accelerates the moving contacts. The rapid separation of the contacts and the use of contact tips composed of a tungsten-carbon alloy mitigate the attendant ablation of contact material.

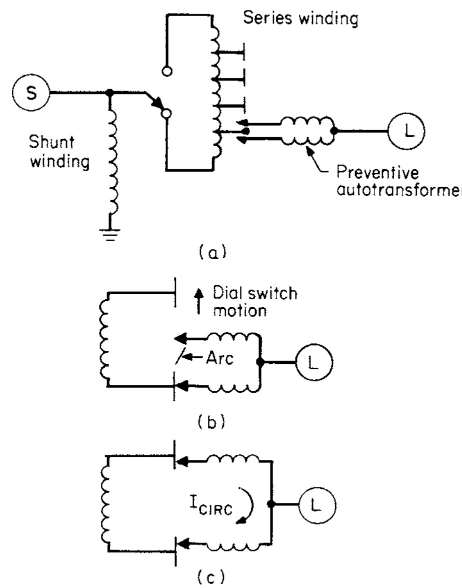


FIGURE 10-100 Operations of the internal mechanisms of the regulator.

A *reversing switch* permits the polarity of the series winding to be reversed relative to the shunt winding, thereby accommodating plus and minus regulation with the same series winding.

A *voltage transformer* and *current transformer* are used to provide voltage and current supplies, respectively necessary for the control to perform its function.

Dielectric protection of the series winding is afforded by the *bypass arrestor.* A surge propagated on the line will be shunted past the regulator. *Lightning arrestors,* often provided at both the source and the load terminals, similarly protect the regulator from overvoltage surge conditions.

Other external appurtenances will include the three bushings (source, load, and SL or common), a *tap position indicator* display of the present operating position of the regulator, and a *control panel enclosure.*

Regulator Control Circuitry. This technology has progressed from a mechanically driven regulating relay beam to highly sophisticated, microprocessor-based digital control circuitry. All regulator controls comprise three major parts:

1. A *voltage-sensing device* which monitors the output voltage of the regulator and sends a signal to the motor drive circuity
2. An *amplifying or switching section,* with or without time delay, which delays and/or transmits the signal
3. A *motor drive circuit* which responds to the signal by closing relay contacts or actuating electronic switches that cause the tap-changing motor to operate the tap changer correct the voltage

The important functions of the typical regulator control are shown in the block diagram in Fig. 10-101.

An *auxiliary voltage source,* such as a nominal 120-V tertiary winding on the main coil and/or a separate voltage transformer, will be provided to supply power to the tap changer motor and the base voltage for deriving the regulator output voltage. This 120 V is scaled down in a *sensing trans-*

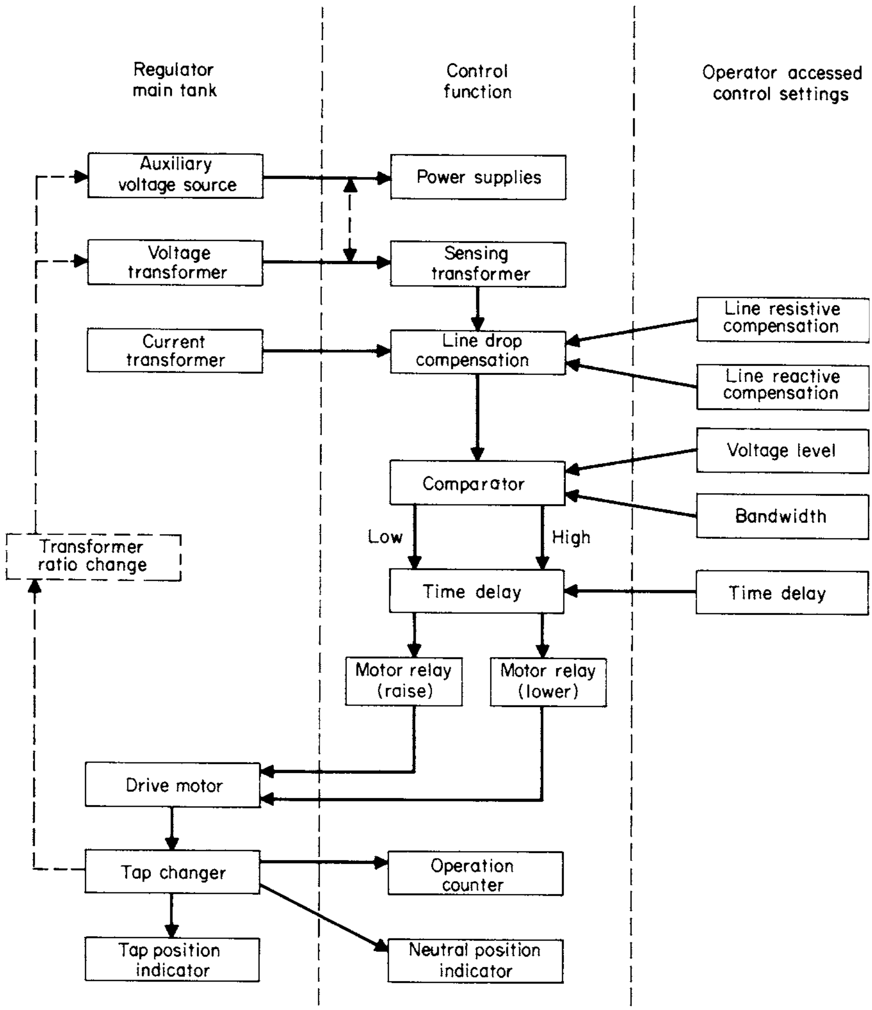


FIGURE 10-101 Regulator control functions.

former which may itself be tapped to facilitate ratio adjustment if the voltage source is not exactly as required. The output of the sensing transformer will, by operator selection, be altered to reflect the voltage drop between the regulator and the load. This drop, which is a function of the load current and the line impedance, is modeled by knowing the current from a *current transformer* and established *line resistive compensation* and *line reactive compensation* parameters.

The voltage, now compensated for line drop, is related in a *comparator* to the desired voltage level as established by a *voltage-level* setting and a *bandwidth* or tolerable voltage range limits setting.

If the compensated voltage exceeds the high limit (or is less than the low limit), a *time delay* is activated. Requiring the voltage to remain out-of-band for a period of typically 30 to 60 s assures that the voltage-level change is of sufficient duration to warrant regulator action.

After the preset delay is satisfied, a *drive motor* is powered and runs the *tap changer* to the next position. Tap-changer motion advances the *operation counter,* moves the *tap position indicator* pointer, and causes illumination of a special *neutral position indicator* light, if in neutral position.

Of course, the new tap position in turn causes a *transformer ratio change,* altering the input to the sensing transformer and closing the control function loop.

10.4.2 Application of Regulators

The rating of single-phase step-voltage regulators is a simple function of the rated range of regulation and the current. For example, a 100-A regulator with a ±10% range of regulation for a 7620-V system will be rated

$$100 \text{ A} \times 0.10 \times 7620 \text{ V} = 76.2 \text{ kVA}$$

Early recognition that this is 10% of circuit kVA will avoid later difficulties.

The ampere rating of regulators can be increased by limiting the range of regulation to accommodate higher line currents. By limiting the range of regulation, the portion of series winding through which line current passes is decreased. This decrease in the portion of series winding "seeing" load current decreases the internal regulator losses. By decreasing losses, heating inside the regulator is decreased. Consequently, the regulator may carry more load current for the same temperature rise. For example, a 100-A regulator, ±10%, may be operated at ±5% regulation at 160 A. Limit-switched taps and overload capabilities are, by ANSI standards, 8 $^3/_4$% (110%), 7$^1/_2$% (120%), 6$^1/_4$% (135%), 5% (160%), to a limit of 668 A.

Oil-immersed step-type regulators, by ANSI standards, are rated at 55°C average winding temperature rise. A 65°C hot-spot winding temperature is also specified in regulator standards. The short-circuit withstand capacity of standard regulators is 25 times normal full-load current for 2 s.

Using single-phase regulators in three-phase installations is a very common application. Configuring the regulators in wye, closed delta, and open delta are acceptable alternatives, depending on system conditions.

The wye connection is invariably used on grounded-neutral 4-wire systems. Regulator sizing is a straightforward adaptation of single-phase principles.

The use of three single-phase regulators in closed delta requires that recognition be made of how current will flow in the regulator. Especially important is the fact that system line current and regulator series winding current are not the same. The application will be specified based on load current, but the current in the regulator series winding will establish the regulator size required. Also, the closed-delta application results in a ±15% range of regulation in the line voltage for a ±10% range section on the individual single-phase regulators. Another point which is significant when modeling for line-drop compensation or sensing a power reversal is to note that the current and voltage signals from the regulators will be displaced by 30° (lead or lag) at unity power factor of the load.

An open-delta connection is sometimes used to save expense. This system is like a wye connection in that line current and series winding current are identical, but like a delta in that the 30° phase shift occurs; in fact, in this case the shift is leading in one regulator and lagging in the other.

The various regulator connections utilize the installed kVA of regulators with differing efficiency. For each 1000 kVA of system capacity, the use of ±10% regulators will require total regulator capacity of as given in the following table:

Grounded wye	100 kVA
Closed delta	123 kVA
Open delta	115 kVA

Another important consideration when applying single-phase regulators in three-phase installations is that certain single-phase regulator connections may be unsafe. In deciding on the proper and safe connection for a given application, three basic phenomena should be considered (1) third harmonics, (2) system line surges, and (3) line faults.

For instance, when single-phase regulators are grounded on an ungrounded system, a resonant circuit is possible between the third-harmonic magnetizing reactance of the regulators and line capacitance. This resonance can intensify third-harmonic voltages to unsafe levels. A similar situation can occur when regulators are ungrounded on either a grounded or an ungrounded system. To protect from line surges, regulator connection is unimportant as long as a bypass arrester and line-to-ground arresters are used. Line-to-ground fault problems may be serious when regulators are ungrounded on a grounded system or grounded on an ungrounded system. A problem may also occur when regulators are connected in either open- or closed-delta configuration on a grounded system. Table 10-11 summarizes safe and unsafe single-phase regulator connections. A feature of regulators is that they may be bypassed in service such that load interruption is not necessary during installation procedures. It is critical that proper switching procedures be followed to avoid the extremely high circulating current which will appear in the series winding if bypassing occurs at other than the neutral tap position. When placing a regulator on the line (see Fig. 10-101), three switches are commonly used: a source switch, load switch, and bypass switch. To place the regulator in service, first the source switch is closed. The regulator is checked out by running the tap changer in the raise and lower directions. The regulator is returned to the neutral position, and the load switch is closed. Caution must be taken to make sure that the regulator will not make an automatic tap change when the load switch is closed. After the load switch is closed, the bypass switch may be opened. To take the regulator out of service, the procedure is reversed. First the regulator is run to the neutral position, the bypass switch is closed, the load switch is opened, and then finally the source switch is opened. Once again, caution must be taken to prevent the regulator from making an automatic tap change during this switching procedure by assuring isolation of the control power.

Regulators may be paralleled if and only if sufficient loop impedance exists to limit circulating currents, based on the maximum difference in the voltages of the two circuits V_1 and V_2. Equally important, the percent impedances of the two circuits must be equal or very close. The common practice is that the two impedances must be close enough so that circulating current does not exceed 10% of rated current when operating on equal voltage taps. For impedance Z_1 of circuit 1 and Z_2 of circuit 2, percent circulating current is given approximately by

$$\%\ \text{circulating current} = \frac{Z_1 - Z_2}{Z_1 + Z_2} \times 100$$

When regulators are paralleled, it is necessary that the tap-changing mechanisms be on as nearly the same tap position as possible. If they are not, a circulating current of magnitude $I_c = (V_1 - V_2)/(Z_1 + Z_2)$ will flow in the loop. The most widely used method for paralleling regulators is the "current-balance" method. Circulating current is separated from load current by means of auxiliary current transformers. This current is fed into the voltage reference circuit to cause the unit to change taps to reduce the circulating current.

Regulators are often applied in series, or cascaded, on the same feeder. In this application, two or more regulators operate to control the voltage profile along with a distribution line. The most important

TABLE 10-11 Safe and Unsafe Regulator Connections

Regulator connection	System connection	Effect: Third harmonic	Effect: Line surges	Effect: Line ground	Conclusion
Grounded Y	Grounded	S	S_1	S	S
Ungrounded Y	Grounded	U	S_1	U	U
Delta	Grounded	S	S_1	S_2	S_2
Open delta	Grounded	S	S_1	S_2	S_2
Grounded Y	Ungrounded	U	S_1	U	U
Ungrounded Y	Ungrounded	U	S_1	S	U
Delta	Ungrounded	S	S_1	S	S
Open delta	Ungrounded	S	S_1	S	S

Key: S = Safe; S_1 = safe if suitable bypass series winding protection is supplied; S_2 = conditionally safe—overexcitation of regulators may lead to their failure if fault allowed to persist; U = unsafe.

consideration in such applications is the time-delay settings on the regulator controls. The modern solid-state control is adjustable from at least 15 to 105 s. When regulators are cascaded, the first regulator (at or closest to the substation) should be set with the shortest time delay, with progressively longer delays farther away from the station. If the settings are reversed, the farthest regulator would attempt to correct all voltage fluctuations first. As the load change appeared at the other regulators on the line, each would attempt to compensate for it in order. When finally the substation unit reacted, it would raise the overall voltage level, making it necessary for each successive unit down the line to back down accordingly.

Setting of the regulator control must be accomplished with care to assure proper regulator operation and proper supply voltage for the users.

The voltage-level setting is the voltage which the regulator is to maintain at its output, expressed on a 120- V base. (Note that this will be the voltage at the "load center" if line-drop compensation is used, as explained later.)

To avoid a hunting condition of the regulator, a *bandwidth* is set to define the limits of acceptable voltage about the voltage-level setting. The stated bandwidth is the total range, such that a regulator set for 120 V with a 2-V bandwidth will be "in-band" if the output is in the range of 119 to 121 V. The bandwidth setting must be larger than the voltage change expected from a single tap change or a hunting situation will occur. Beyond this, it is a qualitative judgment; a lower setting will maintain a closer output voltage tolerance, a higher setting will reduce tap-changer operations, extending the regulator life.

The *time delay* is the time duration outside of the prescribed band required before tap-changer actuation. As noted earlier, this is very important in cascaded operations. Otherwise, it will be set at typically 30 to 60 s to avoid unduly quick responses to line-voltage fluctuations.

The use of *line-drop compensation* will cause the regulator to hold the voltage-level setting at a point remote from the regulator, rather than at the regulator location. The classic illustration of the application of line-drop compensation involves a "load center" some miles from the substation. It is required to hold a given voltage at the load. Given that the line is inductive in nature, this implies holding a higher voltage at the substation, the incremental voltage increase being a function of the line impedance (resistive and reactive) and the line current. Thus, the two line-drop compensation settings are the resistive and reactive models of the line, calibrated in a 0- to 24-V basis, reflecting the drop to be anticipated (on a 120-V base) between the regulator and the load when the system is carrying rated regulator current.

Tables are provided with the regulators to use in determining the proper settings. A simplified example demonstrates the procedure:

1. The wye-connected regulators are rated 76.2 kVA, 7620 V (= 100 A).
2. The load center is 4 mi from the regulators.
3. The feeder to the load center is 2/0 ACSR on 36 in center spacing.

The solution to this example is as follows:

1. Tables provided with the regulator will show
   - ***a.*** Line resistance $\simeq$ 0.90 Ω/mi
   - ***b.*** Line reactance $\simeq$ 0.77 Ω/mi
2. Calculate compensation as
   - ***a.*** $$\frac{\text{CT primary rating}}{\text{Voltage transformer ratio}} \times R/\text{mi} \times \text{miles} = R \text{ comp set}$$
   - ***b.*** $$\frac{\text{CT primary rating}}{\text{Voltage transformer ratio}} \times X/\text{mi} \times \text{miles} = X \text{ comp set}$$

$$R_{\text{comp set}} = 5.7 \text{ or } 6 \text{ V}$$

$$X_{\text{comp set}} = 4.9 \text{ or } 5 \text{ V}$$

Thus the required 120 V will be held at the load. Changes in load current are automatically compensated.

Each line-drop compensation application needs individual consideration. Conditions which would invalidate this example are single-phase or delta connection or the regulators. Also, the example of a precise load center is seldom encountered in practice. For this reason, the setting of line-drop compensation must often be tempered with knowledge of actual system conditions.

Several control accessories are available to provide more sophisticated feeder loading control. These devices have served to make modern feeder regulators the "nerve center" of the distribution system. A *voltage-limit control* device provides automatic limit of regulator output voltage. Settings for both upper and lower limits protect against extremely heavy, light, or unusual loading conditions. A review of the use of line-drop compensation may show that at the highest anticipated loading, the voltage at the regulator is too high for proper customer utilization. Then an upper-limit voltage-limit control may be required to protect a load close to the regulator by overriding the line-drop compensation function. A *reverse power flow detector* can monitor source-side voltage with an internal source-side voltage supply, detect a reversal of power flow direction, and "turn the regulator around" electrically so that it regulates in the proper direction. A *voltage-reduction control,* which can be operated locally or from a remote control, provides automatic voltage reduction in preselected percentages. This is particularly useful where system capacity is close to peak load; studies have shown that a 5% voltage reduction can reduce system load by almost 5%.

10.4.3 Regulator Developments

Innovation associated with step-voltage regulators has concentrated in recent years on the electronic control. The very nature of digital controls now offered routinely with new step-voltage regulators facilitates the mathematical manipulation of the measured line voltage and current into additional system parameters of interest to the user. Thus, controls are available which will display voltage, current, power factor, kW, kvar, various time-integrated demand quantities, and other conditions of interest.

An additional feature available on many controls is the ability to serially communicate this information to a remote-terminal unit (RTU) so that the regulator control becomes the sensory apparatus of the supervisory control and data acquisition (SCADA) system, eliminating the need for multiple transducers and the attendant analog signals.

10.5 POWER CAPACITORS

By JEFFREY H. NELSON and ROBERT L. KLEEB

Definitions of terms used in this section can be found in the IEEE standards and application guides referenced in this section and/or in the *IEEE Authoritative Dictionary of Terms.*

10.5.1 System Benefits of Power Capacitors

Power capacitors provide several benefits to power systems. Among these include power factor correction, system voltage support, increased system capacity, reduction of power system losses, reactive power support, and power oscillation damping.

Power Factor Correction. In general, the efficiency of power generation, transmission, and distribution equipment is improved when it is operated near unity power factor. The least expensive way to achieve near unity power factor is with the application of capacitors. Capacitors provide a static source of leading reactive current and can be installed close to the load. Thus, the maximum efficiency may be realized by reducing the magnetizing (lagging) current requirements throughout the system. Table 10-1 is a simple tool that can be used to determine the kilovars (kvar) required for correcting the system power factor. From Table 10-12, find the row that corresponds to the existing power factor and the column that corresponds to the desired power factor and select the kilowatt multiplier where these intersect. Then simply multiply this factor by the power system load in kilowatts to determine the kilovars required to be installed to achieve the desired power factor.

For example, with a load of 200 kW at 77% power factor, how many capacitor kilovars are needed to correct to a power factor of 95%? At the point where the row for the 77% existing power factor and the column for the desired power factor of 95% intersect, we find a kilowatt multiplier of 0.5. Therefore, the following calculation can be made to determine the kvar required to achieve 95% power factor.

Example

$$\begin{aligned} 3\text{Ø kvar} &= 3\text{Ø load (kW)} \times \text{kilowatt multiplier} \\ &= 200\text{ kW} \times 0.5 \\ &= 100\text{ kvar} \end{aligned} \tag{10-76}$$

System Voltage Support. Power systems are predominately inductive in nature and during peak load conditions or during system contingencies there can be a significant voltage drop between the voltage source and the load. Application of capacitors to a power system results in a voltage increase back to the voltage source, and also past the application point of the capacitors in a radial system. The actual percentage increase of the system voltage is dependent upon the inductive reactance of the system at the point of application of the capacitors. The short-circuit impedance at that point is approximately the same as the inductive reactance; therefore, the 3-phase short-circuit current at that location can be used to determine the approximate voltage rise. The following rule-of-thumb equation is commonly used

$$\Delta V \approx \frac{\text{kvar}_\text{C}}{\text{kVA}_\text{SC}} \times 100\% \tag{10-77}$$

where ΔV = percent voltage rise at the point of the capacitor installation, kvar_C = capacitor 3-phase kvar, kVA_SC= system 3-phase short-circuit kVA at the point of the capacitor installation.

$$\text{kVA}_\text{SC} = \sqrt{3} \times V_{LL} \times I_{SC} \tag{10-78}$$

where V_{LL} = system line-to-line (phase-to-phase) voltage, I_{SC} = is the 3-phase short-circuit current at the point of the capacitor installation.

Increased System Capacity. The application of shunt or series capacitors can affect the power system capacity.

Application of shunt capacitors reduces the inductive reactive current on the power system, and thus reduces the system kVA loading. This can have the effect of increasing system to serve additional load.

Series capacitors are typically used to increase the power carrying capability of transmission lines. Series capacitors insert a voltage in series with the transmission line that is opposite in polarity to the voltage drop across the line, which decreases the apparent reactance and increases the power transfer capability of the line.

Power System Loss Reduction. The installation of capacitors can reduce the current flow in a power system. Since losses are proportional to the square of the current, a reduction in current will lead to reduced system losses.

Reactive Power Support. Capacitors can help support steady-state stability limits and reactive power requirements at generators.

Power Oscillation Damping. Controlled series capacitors can provide an active damping for power oscillations that many large power systems experience. They can also provide support after significant disturbances to the power system and allow the system to remain in synchronous operation.

TABLE 10-12 Power Factor Correction Kilowatt Multiples

Present power factor percentage	Corrected power-factor percentage																				
	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
50	0.982	1.008	1.034	1.060	1.086	1.112	1.139	1.165	1.192	1.220	1.248	1.276	1.306	1.337	1.369	1.403	1.442	1.481	1.529	1.590	1.732
51	0.937	0.962	0.989	1.015	1.041	1.067	1.094	1.120	1.147	1.175	1.203	1.231	1.261	1.292	1.324	1.358	1.395	1.436	1.484	1.544	1.687
52	0.893	0.919	0.945	0.971	0.997	1.023	1.050	1.076	1.103	1.131	1.159	1.187	1.217	1.248	1.280	1.314	1.351	1.392	1.440	1.500	1.643
53	0.850	0.876	0.902	0.928	0.954	0.980	1.007	1.033	1.060	1.088	1.116	1.144	1.174	1.205	1.237	1.271	1.308	1.349	1.397	1.457	1.600
54	0.809	0.835	0.861	0.887	0.913	0.939	0.966	0.992	1.019	1.047	1.075	1.103	1.133	1.164	1.196	1.230	1.267	1.308	1.356	1.416	1.559
55	0.769	0.795	0.821	0.847	0.873	0.899	0.926	0.952	0.979	1.007	1.035	1.063	1.090	1.124	1.156	1.190	1.228	1.268	1.316	1.377	1.519
56	0.730	0.756	0.782	0.808	0.845	0.860	0.887	0.913	0.940	0.968	0.996	1.024	1.051	1.085	1.117	1.151	1.189	1.229	1.277	1.338	1.480
57	0.692	0.718	0.744	0.770	0.796	0.822	0.849	0.875	0.902	0.930	0.958	0.986	1.013	1.047	1.079	1.113	1.151	1.191	1.239	1.300	1.442
58	0.655	0.681	0.707	0.733	0.759	0.785	0.812	0.838	0.865	0.893	0.921	0.949	0.976	1.010	1.042	1.076	1.114	1.154	1.202	1.263	1.405
59	0.618	0.644	0.670	0.696	0.722	0.748	0.775	0.801	0.828	0.856	0.884	0.912	0.939	0.973	1.005	1.039	1.077	1.117	1.165	1.226	1.368
60	0.584	0.610	0.636	0.662	0.688	0.714	0.741	0.767	0.794	0.822	0.850	0.878	0.905	0.939	0.971	1.005	1.043	1.083	1.131	1.192	1.334
61	0.549	0.575	0.601	0.627	0.653	0.679	0.706	0.732	0.759	0.787	0.815	0.843	0.870	0.904	0.936	0.970	1.008	1.048	1.096	1.157	1.299
62	0.515	0.541	0.567	0.593	0.619	0.645	0.672	0.698	0.725	0.753	0.781	0.809	0.836	0.870	0.902	0.936	0.974	1.014	1.062	1.123	1.265
63	0.483	0.509	0.535	0.561	0.587	0.613	0.640	0.666	0.693	0.721	0.749	0.777	0.804	0.838	0.870	0.904	0.942	0.982	1.030	1.091	1.233
64	0.450	0.476	0.502	0.528	0.554	0.580	0.607	0.633	0.660	0.688	0.716	0.744	0.771	0.805	0.837	0.871	0.909	0.949	0.997	1.058	1.200
65	0.419	0.445	0.471	0.497	0.523	0.549	0.576	0.602	0.629	0.657	0.685	0.713	0.740	0.774	0.806	0.840	0.878	0.918	0.966	1.027	1.169
66	0.388	0.414	0.440	0.466	0.492	0.518	0.545	0.571	0.598	0.626	0.554	0.682	0.709	0.743	0.775	0.809	0.847	0.887	0.935	0.996	1.138
67	0.358	0.384	0.410	0.436	0.462	0.488	0.515	0.541	0.568	0.596	0.624	0.652	0.679	0.713	0.745	0.779	0.817	0.857	0.905	0.966	1.108
68	0.329	0.355	0.381	0.407	0.433	0.459	0.486	0.512	0.539	0.567	0.595	0.623	0.650	0.684	0.716	0.750	0.788	0.828	0.876	0.937	1.079
69	0.299	0.325	0.351	0.377	0.403	0.429	0.456	0.482	0.509	0.537	0.565	0.593	0.620	0.654	0.866	0.720	0.758	0.798	0.840	0.907	1.049
70	0.270	0.296	0.322	0.348	0.374	0.400	0.427	0.453	0.480	0.508	0.536	0.564	0.591	0.625	0.657	0.691	0.729	0.769	0.811	0.878	1.020
71	0.242	0.268	0.294	0.320	0.346	0.372	0.399	0.425	0.452	0.480	0.508	0.536	0.563	0.597	0.629	0.663	0.701	0.741	0.783	0.850	0.992
72	0.213	0.239	0.265	0.291	0.317	0.343	0.370	0.396	0.423	0.451	0.479	0.507	0.534	0.568	0.600	0.634	0.672	0.712	0.754	0.821	0.963
73	0.186	0.212	0.238	0.264	0.290	0.316	0.343	0.369	0.396	0.424	0.452	0.480	0.507	0.541	0.573	0.607	0.645	0.685	0.727	0.794	0.936
74	0.159	0.185	0.211	0.237	0.263	0.289	0.316	0.342	0.369	0.397	0.425	0.453	0.480	0.514	0.546	0.580	0.618	0.658	0.700	0.767	0.909
75	0.132	0.158	0.184	0.210	0.236	0.262	0.289	0.315	0.342	0.370	0.398	0.426	0.453	0.487	0.519	0.553	0.591	0.631	0.673	0.740	0.882
76	0.105	0.131	0.157	0.183	0.209	0.235	0.262	0.288	0.315	0.343	0.371	0.399	0.426	0.460	0.492	0.526	0.564	0.604	0.652	0.713	0.885
77	0.079	0.105	0.131	0.157	0.183	0.209	0.236	0.262	0.289	0.317	0.345	0.373	0.400	0.434	0.466	0.500	0.538	0.578	0.620	0.687	0.829
78	0.053	0.079	0.105	0.131	0.157	0.182	0.210	0.236	0.263	0.291	0.319	0.347	0.374	0.408	0.440	0.474	0.512	0.552	0.594	0.661	0.803
79	0.026	0.052	0.078	0.104	0.130	0.156	0.183	0.209	0.236	0.264	0.292	0.320	0.347	0.381	0.413	0.447	0.485	0.525	0.567	0.634	0.776

80	0.000	0.026	0.052	0.078	0.104	0.130	0.157	0.183	0.210	0.238	0.266	0.294	0.321	0.355	0.387	0.421	0.459	0.499	0.541	0.608	0.750
81		0.000	0.026	0.052	0.078	0.104	0.131	0.157	0.184	0.212	0.240	0.268	0.295	0.329	0.361	0.395	0.433	0.473	0.515	0.582	0.724
82			0.000	0.026	0.052	0.078	0.105	0.131	0.158	0.186	0.214	0.242	0.269	0.303	0.335	0.369	0.407	0.447	0.489	0.556	0.698
83				0.000	0.026	0.052	0.079	0.105	0.132	0.160	0.188	0.216	0.243	0.277	0.309	0.343	0.381	0.421	0.463	0.530	0.672
84					0.000	0.026	0.053	0.079	0.106	0.134	0.162	0.190	0.217	0.251	0.283	0.317	0.355	0.395	0.437	0.504	0.645
85						0.000	0.027	0.053	0.080	0.108	0.136	0.164	0.191	0.225	0.257	0.291	0.329	0.369	0.417	0.478	0.620
86								0.026	0.053	0.081	0.109	0.137	0.167	0.198	0.230	0.265	0.301	0.343	0.390	0.451	0.593
87									0.027	0.055	0.082	0.111	0.141	0.172	0.204	0.238	0.275	0.317	0.364	0.425	0.567
88										0.028	0.056	0.084	0.114	0.145	0.177	0.211	0.248	0.290	0.337	0.398	0.540
89											0.028	0.056	0.086	0.117	0.149	0.183	0.220	0.262	0.309	0.370	0.512
90												0.028	0.058	0.089	0.121	0.155	0.192	0.234	0.281	0.342	0.484
91													0.030	0.061	0.093	0.127	0.164	0.206	0.253	0.314	0.456
92														0.031	0.063	0.097	0.134	0.176	0.223	0.284	0.426
93															0.032	0.066	0.103	0.145	0.192	0.253	0.395
94																0.034	0.071	0.113	0.160	0.221	0.363
95																	0.037	0.079	0.126	0.187	0.328
96																		0.042	0.089	0.150	0.292
97																			0.047	0.108	0.251
98																				0.061	0.203
99																					0.142

10.5.2 Capacitor Units

The terms capacitor unit, power capacitor or capacitor may be used interchangeably.

The base standard which covers the ratings and requirements for capacitor units is IEEE 18[1]. This standard applies to capacitors rated 216 V or higher, 2.5 kvar or more, and designed for shunt connection on alternating current (ac) power systems operating at a nominal power frequency of 50 or 60 Hz. IEEE 18 covers service conditions, ratings and capabilities, manufacturing, and testing of shunt capacitor units.

IEEE 824[2] specifies ratings, capabilities and testing requirements in addition to IEEE 18[1] for capacitor units to be utilized in series capacitor banks.

IEEE 1531[3] covers considerations to take into account when specifying units to be used in a harmonic filter bank.

There are typically two voltage classes of capacitor units applied and they are divided into the categories of low-voltage capacitors (below 1000 V) and high-voltage capacitors (1000 V and above).

Low-Voltage Capacitor Units. Low-voltage capacitor units are typically applied in industrial power systems. They are generally available in voltage ratings from 240 to 600 V over the range of 2.5 to 100 kvar 3-phase. Low-voltage capacitors are typically dry type, metalized polypropylene film (see Fig. 10-102). Typical voltage and kvar ratings are listed in IEEE 18[1].

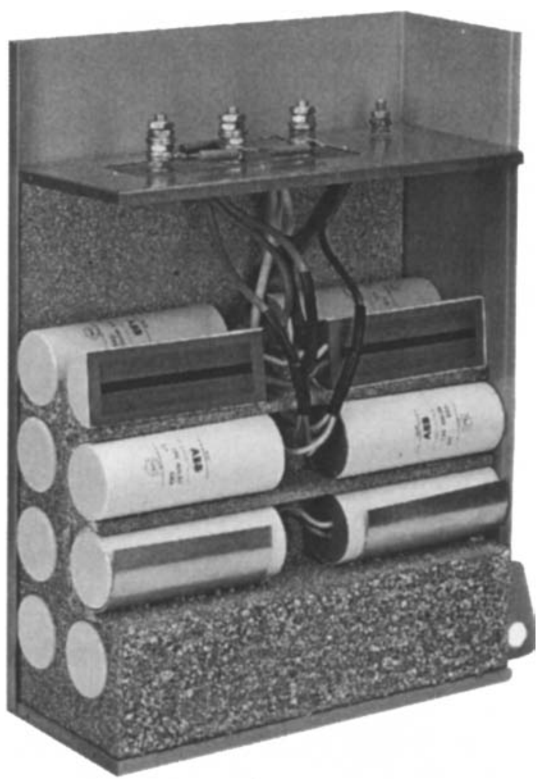


FIGURE 10-102 Typical low-voltage capacitor unit with self-healing metalized polypropylene capacitor elements, liquid-free enclosure and 3-phase connection terminals.

High-Voltage Capacitor Units. High-voltage capacitor units are typically connected to industrial power systems or to utility transmission and distribution power systems. The units are generally available in voltage ratings from 2.4 to 25 kV. Modern manufacturing techniques continue to decrease the size of high-voltage capacitor units which continues to increase the available kvar ratings for units. Typical voltage and kvar ratings are listed in IEEE 18[1].

The capacitor elements in present day high-voltage capacitor units are manufactured using an *all-film dielectric*, typically two layers of polypropylene film between two layers of thin aluminum foil. The capacitor elements are connected in a series/parallel combination to achieve the desired voltage and kvar rating. The elements are placed in a metal enclosure with bushings for external connections and impregnated with a dielectric fluid, usually under vacuum. The unit is then hermetically sealed (see Fig. 10-103).

Capacitor Unit Ratings. IEEE 18[1] establishes the following standard ratings for capacitor units, as applied under the normal service conditions and ambient temperatures defined in the standard:

1. Voltage, rms (terminal-to-terminal)
2. Terminal to case (or ground) insulation class
3. Reactive power
4. Number of phases
5. Frequency

Capacitors are intended to be operated at or below their rated voltage. They are designed to operate at overvoltages of 10% for bank contingencies, such as the failure of an individual unit or element, or system contingencies, such as a high voltage. Note that the overvoltage capability of the individual capacitor elements inside the unit is 110% of its portion of the unit voltage rating. The capacitor unit is capable of operating continuously under contingency conditions provided that the peak voltage

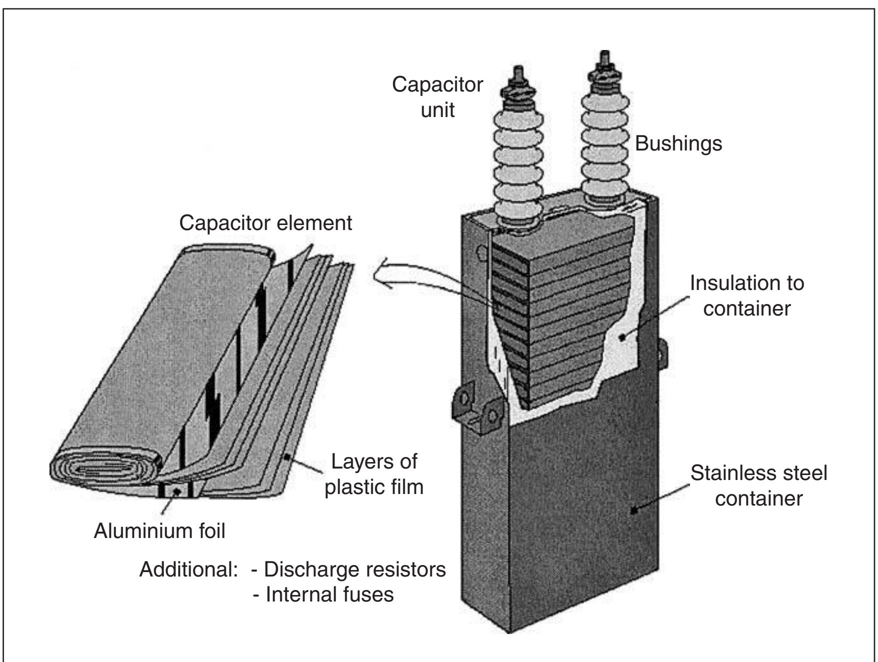


FIGURE 10-103 Typical high-voltage capacitor unit.

does not exceed 120% of rated peak voltage, including harmonics. The maximum peak voltage rating of the capacitor including harmonics, but excluding transients, is equal to

$$V_{pk} = 1.2\sqrt{2}(V_{rms} \text{ rated}) \tag{10-79}$$

Also, the unit is designed to operate continuously if the current does not exceed 135% of nominal rms current based on rated kvar and rated voltage, and the kvar output does not exceed 135% of rated kvar. The total rms current is equal to

$$I_{rms} = \sqrt{(I_1)^2 + (I_2)^2 + (I_3)^2 + \cdots + (I_n)^2} \tag{10-80}$$

where n = harmonic number.

The kvar output of the capacitor increases as the square of the applied voltage

$$\text{kvar}_{E_2} = \text{kvar}_{E_1}\frac{(E_2)^2}{(E_1)^2} \tag{10-81}$$

For example, a 200-kvar, 7200-V capacitor unit will supply 242 kvar at 7920 V

$$\text{kvar}_{E_2} = 200\,\frac{(7920)^2}{(7200)^2}$$

$$\text{kvar}_{E_2} = 242 \text{ kvar}$$

Transient Overvoltage and Overcurrent Withstand Capabilities. When the installation of a switched rack or bank is contemplated, other nearby capacitor equipment must be evaluated.

Capacitor banks switched in a back-to-back configuration can create very high peak magnitude and high frequency current transients.

Also, when switching a capacitor bank without any type of transient suppression a voltage transient is generated, typically in the range of 300 to 800 Hz. This undamped voltage transient can couple through a power transformer to a lower voltage system and be transmitted several miles away. If a capacitor bank is installed on the lower voltage system, a resonant condition may exist. If the resonant frequency of the LC circuit on the high-voltage bus where the capacitor bank is being switched is approximately equal to the LC circuit at the lower voltage capacitor bank, the voltage transient will be magnified. This phenomenon is known as voltage magnification.

These switching conditions could result in high overvoltage and overcurrent transients. The use of series reactors, preinsertion inductors or resistors, and/or special switching devices is sometimes required to reduce these transients to safe levels.

Transient overvoltage and overcurrent capabilities for shunt capacitors are covered in IEEE 1036[4]. Prior to 2002, they were covered in IEEE 18[1].

Capacitance Tolerance. In the 2002 revision to IEEE 18, the allowable capacitance tolerance for shunt capacitor units was changed to 0% from +10% of the nominal value based on rated kvar, voltage and frequency, measured at 25°C uniform case and internal temperature. In previous versions of IEEE 18, the allowable capacitance tolerance was 0% to +15%. However, with today's modern manufacturing practices, capacitance tolerances typically do not exceed +5% for standard units.

A tighter tolerance may be specified for units to be utilized in a harmonic filter capacitor bank, for example, −2% to +2% of the nominal capacitance. This tolerance may be applied to the capacitance of the entire harmonic filter bank instead of the individual units.

Capacitance tolerances for series capacitors are specified on a per phase basis for the entire bank in IEEE 824.

Discharge Resistors. Under certain conditions, when line voltage is removed from a power capacitor, the possibility exists that the unit will retain an extremely high charge even days later. This characteristic of retaining such a charge is demonstrated by the high efficiency and low-loss operation of a power capacitor. To eliminate this hazard, all power capacitors contain internal-discharge resistors. This resistor assembly will reduce the terminal voltage from line voltage to 50 V within 5 min of deenergization for a capacitor rated higher than 600 V ac, and within 1 min for capacitors rated 600 V ac or less.

Capacitor Unit Losses. Present-day high-voltage capacitors operate at a lower watts loss per kvar than do low-voltage capacitors. For example, capacitors that utilize an *all-film dielectric* may operate will losses of less than 0.1 W per kvar. Low-voltage capacitors using *metalized polypropylene dielectric* may experience losses of near 0.5 W per kvar.

The cost of operating high-voltage capacitors is lower per kvar than low-voltage capacitors because of the basic difference in dielectric materials which allows high-voltage capacitor to be operated more efficiently.

Capacitor Tank Rupture. Capacitor tank rupture of high-voltage capacitors will occur if the total energy applied to the capacitor unit under failure conditions is greater than the ability of the capacitor tank to withstand such energy. Tank-rupture curves are essential to the correct selection of fuse links for overcurrent protection of externally fused capacitors. A typical tank rupture curve is shown in Fig. 10-104.

Temperature. Although very efficient, power capacitors do consume some power and generate heat. This heat must be adequately ventilated when enclosed or exposed to higher than normal ambient temperatures.

The maximum operating temperature for power capacitors is specified in IEEE 18 for various mounting arrangements. The minimum operating temperature is $-40°C$.

Capacitor applications must be designed for adequate overvoltage and "corona" capabilities, since the partial discharge (corona) characteristics vary with temperature. It is necessary to consider the full range of temperatures to which the capacitors will be exposed, in both the energized and deenergized modes.

10.5.3 Shunt Capacitors

The most common application of power capacitors is shunt connected capacitors and they are either energized continuously or switched on and off during load cycles.

Common Shunt Capacitor Connections. Figure 10-105 shows four of the most common capacitor connections: 3-phase grounded wye, 3-phase ungrounded wye, 3-phase delta, and single phase. Grounded or ungrounded wye connections are usually made on medium and high-voltage systems, whereas delta and single-phase connections are usually made on low-voltage systems.

Grounded-wye capacitors can bypass some line surges to ground and therefore exhibit a certain degree of self-protection from transient voltages and lightning surges. The grounded-wye connection also provides a low impedance path for harmonics.

If the capacitors are electrically connected ungrounded-wye, the maximum fault current would be limited to 3-times line current. If the available fault current exceeds the tank rupture curve the use of current limiting fuses must be considered.

Low-Voltage Shunt Capacitors. Low-voltage shunt capacitors are typically applied in industrial power systems, in voltage ratings from 240 to 600 V over the range of 2.5 to 100 kvar 3-phase. They can be used for power factor improvement and voltage support on low voltage systems where adjustable speed drives, power electronics, and other industrial equipment are applied.

Low-voltage capacitors made with metalized polypropylene are self-healing. That is, the conductors consist of very thin layers of metal deposited on the film dielectric. In the event of short circuit, the conductor vaporizes to eliminate the fault with negligible loss of capacitance and continued operation. External fuses are not required on these units although some users add them for additional protection of their connections and ground faults.

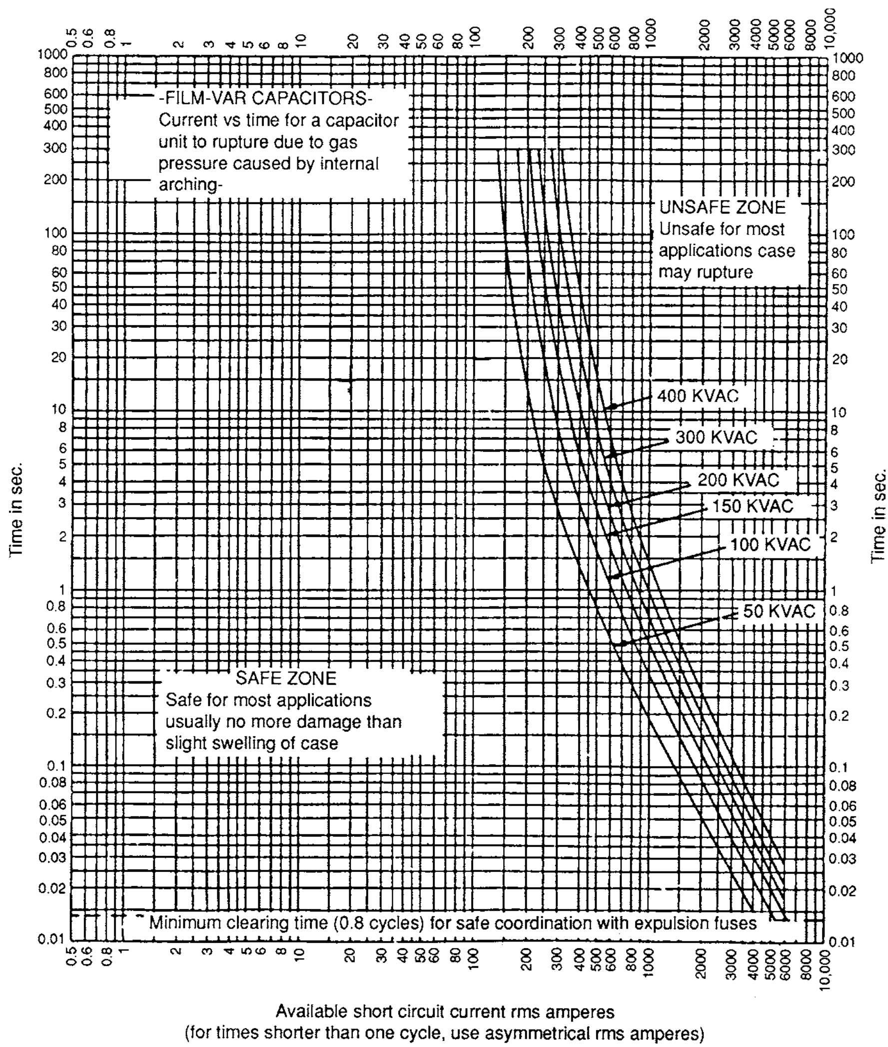


FIGURE 10-104 High-voltage capacitor tank-rupture curve.

Distribution Shunt Capacitors. Capacitors for overhead distribution systems can be pole-mounted typically in banks of 300 to 3600 kvar at nearly any system voltage up to 34.5 kV phase-to-phase. Pad-mounted capacitor equipments are used for underground distribution systems in the same range of sizes and voltage ratings. See Fig. 10-106 for a typical pole-mounted capacitor bank.

The majority of the power capacitor equipment installed on medium voltage distribution systems is connected grounded wye. With the grounded-wye connection, tanks and frames of switching equipment are at ground potential. This provides increased personnel safety. Grounded-wye connections provide for faster fuse operation in case of a capacitor failure.

Distribution shunt capacitors can be protected with individual unit fuses or group fused. Tank rupture curves are essential to the correct selection of fuse links for overcurrent protection of any capacitor installation. Fuse selections should be based upon the coordination of the fuse-link maximum clearing curve (Fig. 10-107) and the high-voltage capacitor tank-rupture curve (Fig. 10-104).

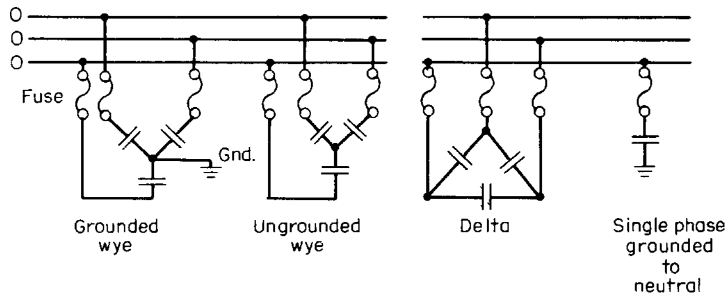


FIGURE 10-105 Common methods for connecting power-factor capacitors.

Several fundamental principles must be observed in the selection of fuses for capacitor applications:

1. The fuse link must be capable of continuously carrying 135% of the nominal capacitor current as a minimum. Higher values may be required when high harmonic currents are present.
2. The fuse cutout must have sufficient interrupting capacity to successfully handle the available fault current, clearing voltage, and available energy before the capacitor tank ruptures.
3. The fuse link must withstand, without damage, the normal transient current during bank energization or de-energization. Similarly, it must withstand the capacitor unit's discharge current during a terminal-to-terminal short.
4. For ungrounded-wye banks, maximum fault current is usually limited to 3 times normal line current. The fuse link must clear within 5 min at 95% of available fault current.

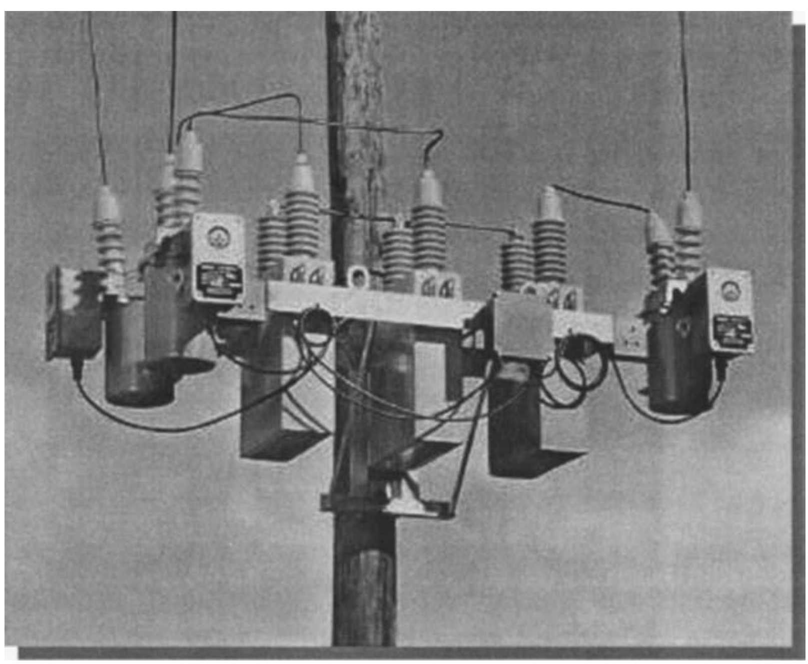


FIGURE 10-106 Typical pole-mounted distribution capacitor bank.

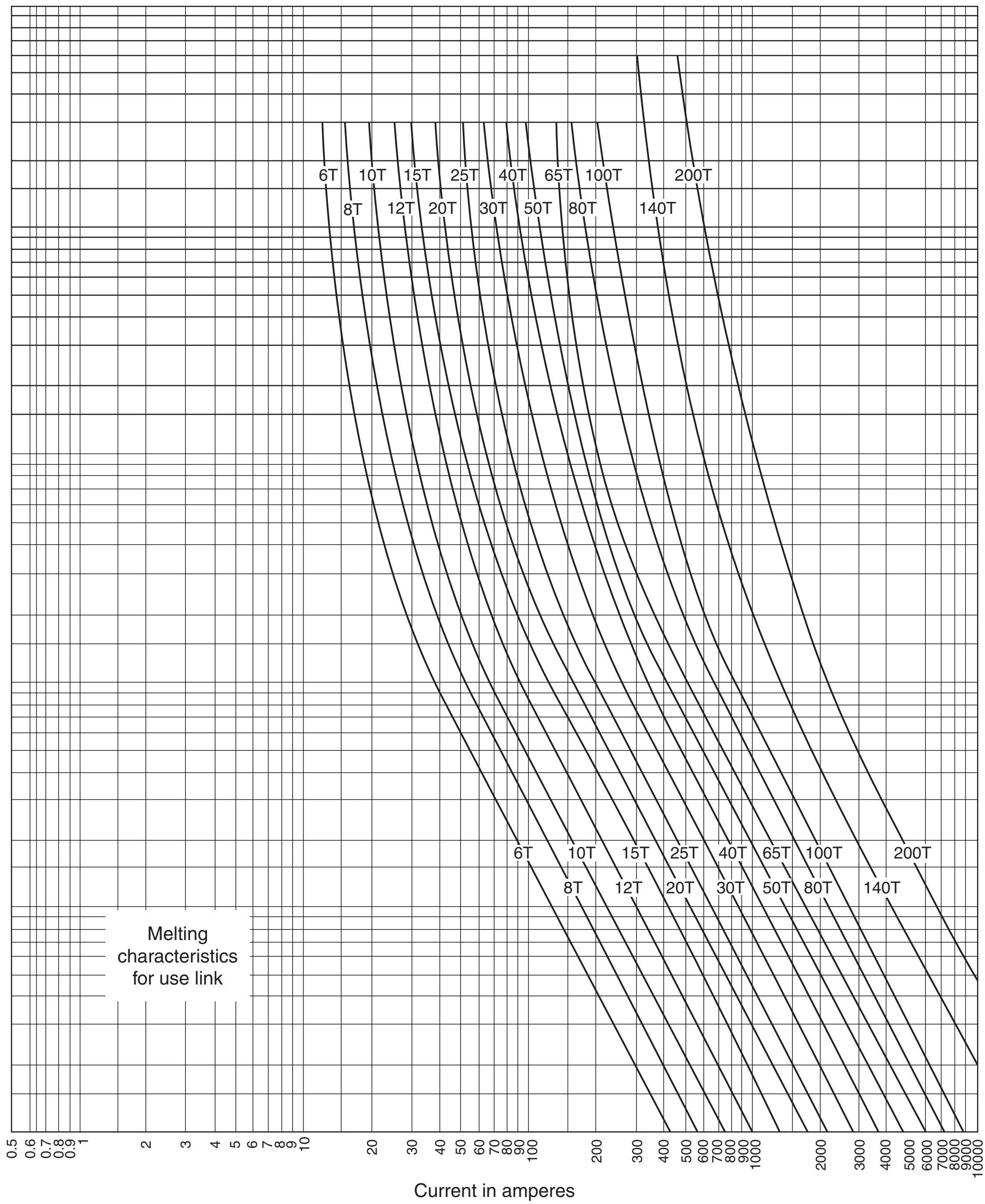


FIGURE 10-107 Typical time-current curve for an expulsion fuse.

5. For effective capacitor protection, maximum asymmetric rms fault current should not exceed the current value at the intercept point of the capacitor tank-rupture time-current characteristic (TCC) curve and the minimum time shown on the fuse maximum-clearing TCC curve.
6. The maximum-clearing TCC curve of the fuse link must coordinate with the tank-rupture TCC curve of the capacitor unit.

For more detailed information on the application of shunt capacitors in distributions systems refer to IEEE 10364[4]. For application guidance on external fuses for shunt capacitors refer to IEEE C37.48[5].

Substation Shunt Capacitor Banks. Substation shunt capacitor banks are usually installed grounded-wye, ungrounded-wye or delta. Delta connected shunt banks are usually at lower voltages, for example, 2400 V. Grounded-wye shunt capacitor bank installations are typically more economical than ungrounded-wye banks because the neutral does not have to be insulated for the system voltage. This advantage increases with system voltage. Also, the transient recovery voltages (TRV) for grounded-wye banks are less than ungrounded-wye, thus reducing the TRV requirements for switching equipment.

A disadvantage of grounded-wye shunt capacitor banks is that there is a possibility of high-frequency switching transients in the ground grid that could induce transients in relay and control cables resulting in erroneous relay operations.

Ungrounded-wye shunt capacitor banks may be applied at any system voltage as long as the insulation and TRV requirements are considered.

For shunt banks that have only one series group, it is more desirable to install an ungrounded wye bank. In a grounded wye bank with one series group, if one unit completely shorts and the phase-to-ground fault current is too high it will rupture the unit. With an ungrounded-wye bank, the fault current is limited to three times the bank current.

There are three common substation shunt capacitor bank types utilized today: externally fused, internally fused, and fuseless. These three types are briefly described in this section.

For more detailed information on the application and protection of substation shunt capacitor banks, refer to IEEE 1036[4], IEEE C37.99[6] and IEEE C37.48[5].

Externally Fused Shunt Capacitor Banks. An externally fused shunt capacitor bank is made up of series groups of parallel connected capacitor units. Each unit is individually fused with an expulsion or current-limiting fuse mounted external to the unit. A typical schematic for an externally fused shunt capacitor bank is shown in Fig. 10-108.

A capacitor unit voltage rating is chosen usually to limit the number of series groups to as few as possible, while achieving the desired voltage rating of the bank. This usually results in the simplest design and highest sensitivity for unbalance protection.

Fewer series groups may mean more units in parallel in each series group. This could require the use of current limiting unit fuses instead of expulsion type fuses. When a unit fails in an externally fused bank the other units in parallel with it will discharge into the failed unit through its individual fuse. When the energy from this discharge exceeds the energy capability of an expulsion fuse, the user will have to specify current limiting fuses which have a higher energy capability. To avoid this, the user can specify a lower voltage rating for the units and design the bank with more series groups and fewer units in parallel. Another option is to split the bank in two sections, called a double-wye configuration. For principles of fuse selection for the individual units, refer to the earlier section on distribution shunt capacitors. For more detailed guidance on fusing refer to IEEE C37.48[5] and IEEE C37.48.1[7].

It is also important to keep a minimum number of units in parallel in each series group. When a fuse operates and removes a unit from a series group the voltage will increase on the remaining units in parallel with it. It is desirable to have enough units in parallel that the 110% overvoltage capability will not be exceeded for the loss of one unit. There are curves available to help determine the voltage on remaining capacitor units in a series group based on the percentage of capacitor units removed from the series group. These curves are located in IEEE 1036[4], which covers the application of shunt capacitors. The curves were previously located in an appendix of IEEE 18[1], but were removed from the 2002 revision.

A typical externally-fused shunt capacitor bank is shown in Fig. 10-109.

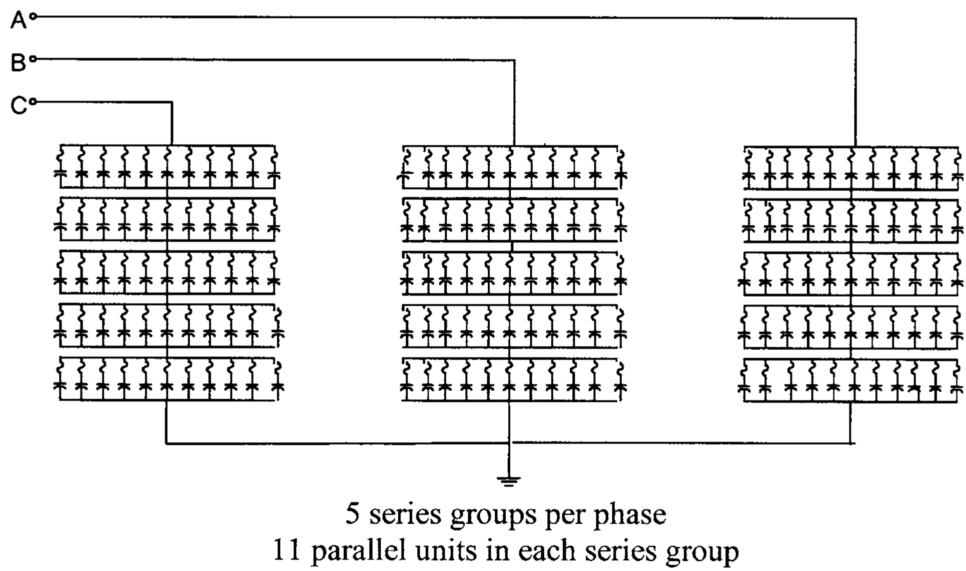


FIGURE 10-108 Schematic of a typical externally fused shunt capacitor bank.

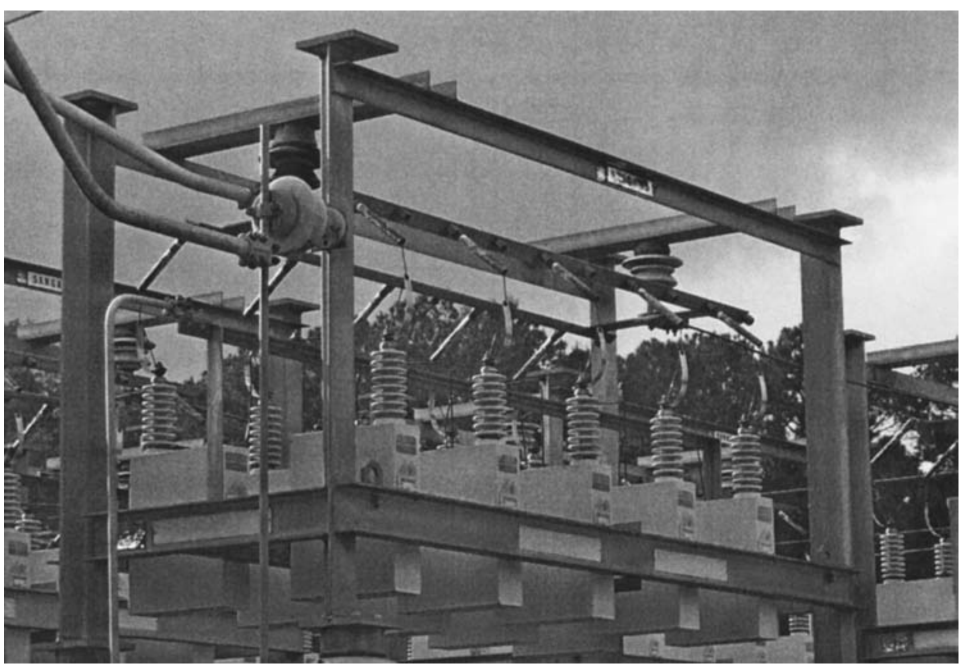


FIGURE 10-109 Typical externally fused shunt capacitor bank.

Internally Fused Shunt Capacitor Banks. An internally fused capacitor bank is made up of series and/or parallel connected capacitor units. The units are made up of capacitor elements connected in series/parallel combinations to achieve the desired rated voltage and kvar rating. A fuse is put in series with each element inside the unit case. An internally fused unit is shown schematically in Fig. 10-110. When a failure of a capacitor element occurs, its individual fuse will operate to isolate the element. The unit typically has a large number of elements in parallel so that the loss of one element does not have a significant increase in voltage on the remaining elements in parallel with the failed element. The kvar rating of internally fused units is typically much larger than externally fused units.

Internally fused capacitor banks typically have fewer units in parallel and more series groups than externally fused capacitor banks. A typical schematic for an internally fused shunt capacitor bank is shown in Fig. 10-111. It is usually desirable to have at least two units

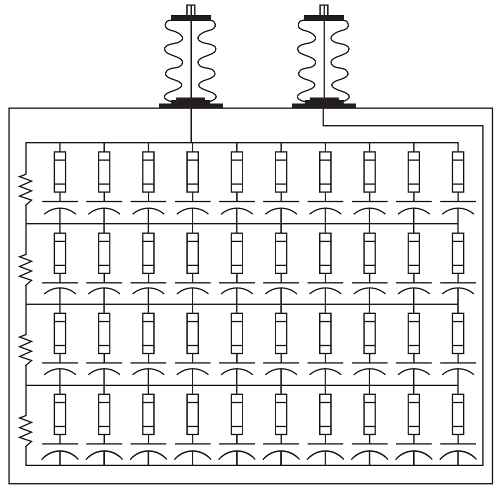


FIGURE 10-110 Internal connection diagram of a typical internally fused capacitor unit.

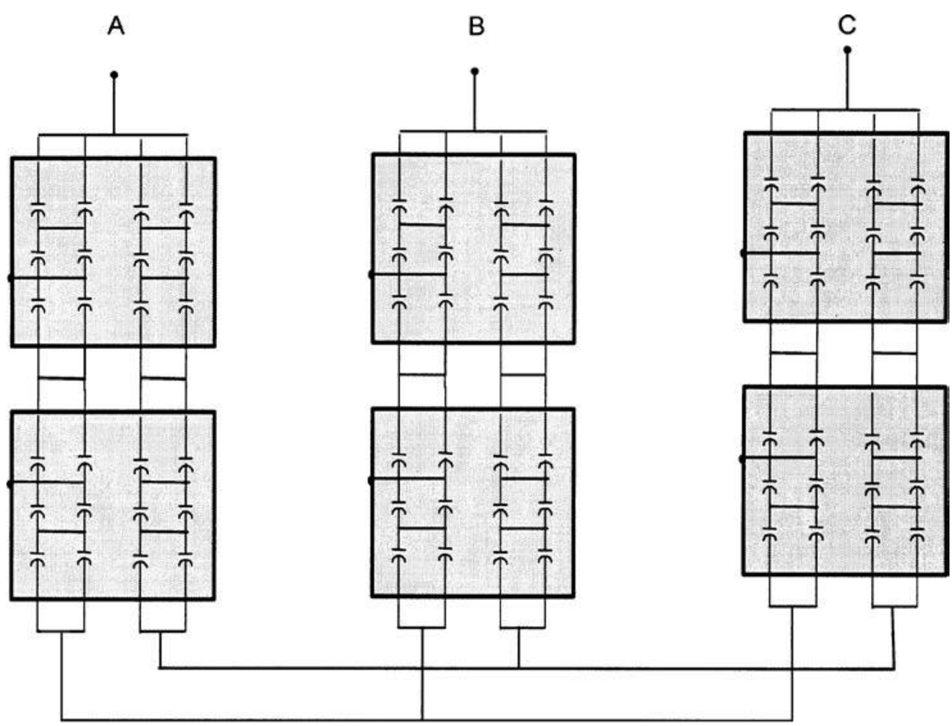


FIGURE 10-111 Schematic of a typical internally fused shunt capacitor bank.

in parallel. If one unit has a large number of failed elements, the unit in parallel will help keep down the terminal-to-terminal voltage. The maximum number of capacitor units that can put in parallel is determined by the energy capability of the internally fused unit.

A typical internally fused capacitor bank is shown in Fig. 10-112.

Fuseless Shunt Capacitor Banks. A fuseless shunt capacitor bank is not simply an externally fused capacitor bank with the fuses removed. It is a different arrangement of capacitor units designed to operate without fuses. A fuseless shunt capacitor bank is made up of units connected in series phase-to-ground or phase-to-neutral. This arrangement of units is called a series string. The unit voltage is selected to achieve the desired voltage rating of the bank. The kvar rating of the unit is chosen to achieve a desired reactive output per each string. Then the bank kvar rating can be increased in these increments by putting multiple series strings of capacitors in parallel. A schematic for a typical fuseless

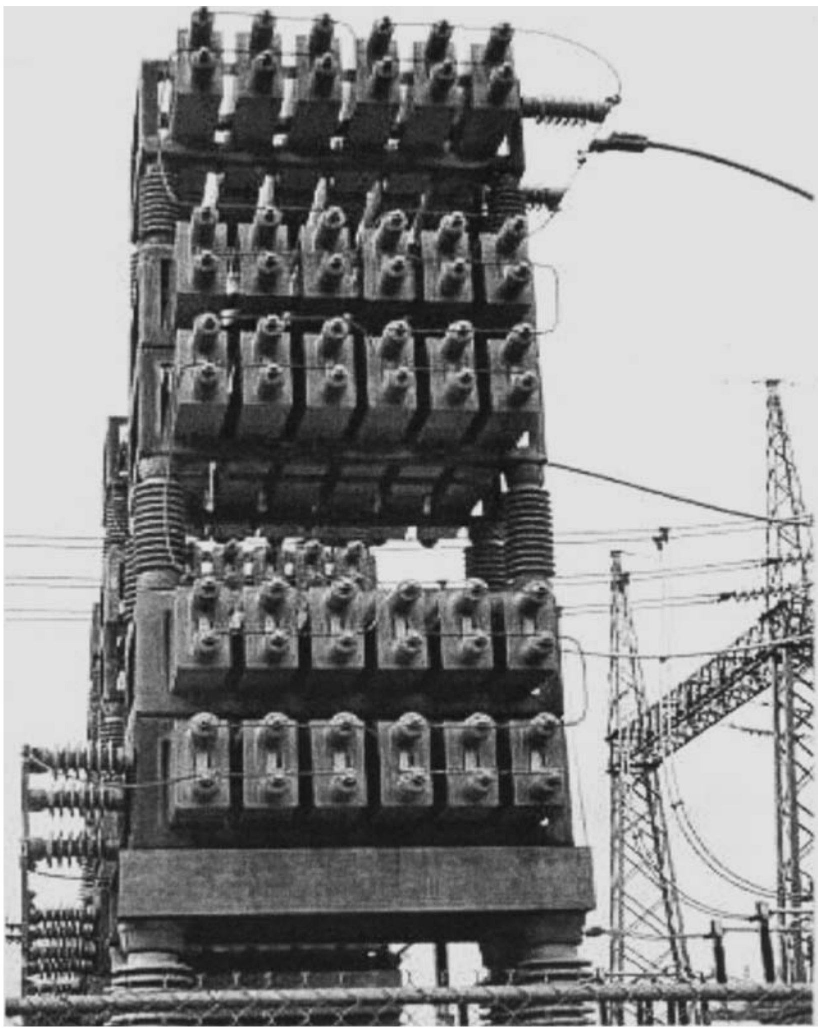


FIGURE 10-112 Example of an internally fused shunt capacitor bank.

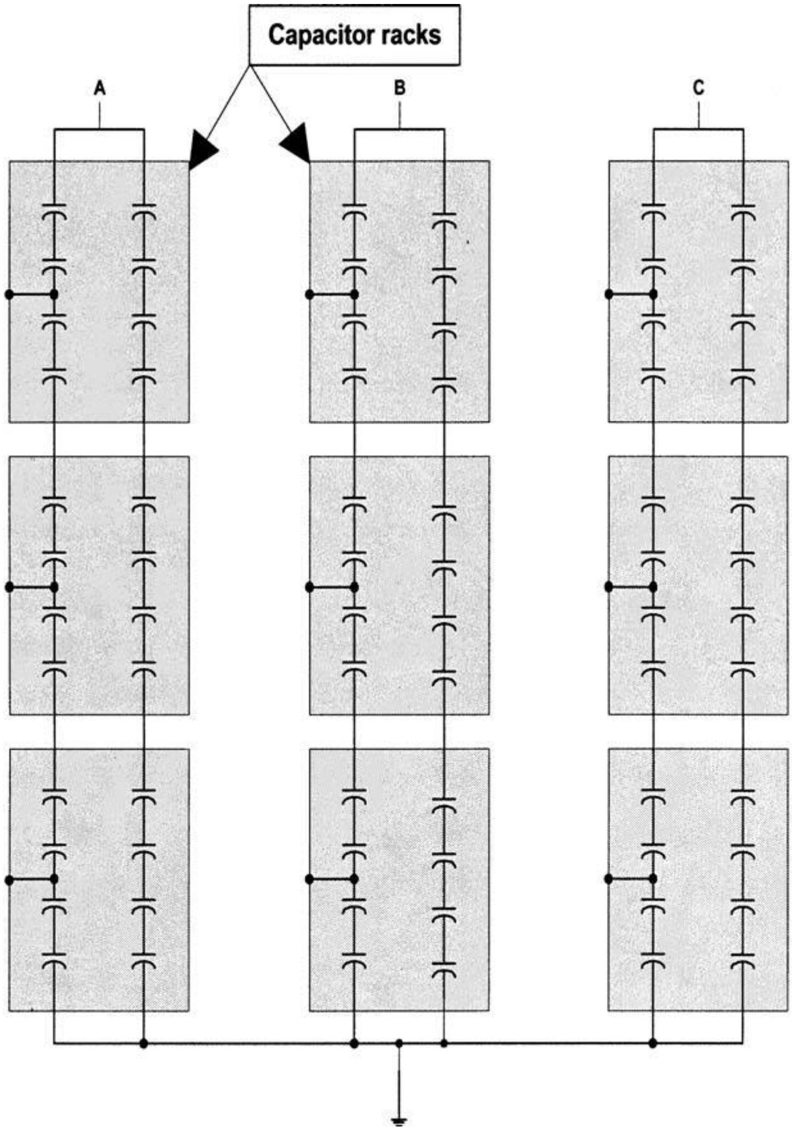


FIGURE 10-113 Schematic of a typical fuseless shunt capacitor bank (Two parallel series strings of twelve units in series.)

shunt capacitor bank is shown in Fig. 10-113. The fuseless capacitor bank is made up of parallel groups of series units, while the externally fused bank is made up of series groups of parallel units.

Fuseless capacitor bank installations were made possible with the design of the *all-film dielectric* capacitor unit, described earlier in the Sec. 10.5.2. The common failure mode of an element in an all-film dielectric unit is for it to fail shorted. When it shorts, the aluminum foil melts and welds together creating a low resistance point that will continue to carry the current through the capacitor indefinitely.

Typically, kvar ratings of units in fuseless capacitor banks are larger than units used in externally fused capacitor banks. This is because fuseless banks do not have the same design consideration of keeping a minimum number of units in parallel as discussed previously for externally fused banks. With the units connected in series, if one element in a unit fails, the voltage will increase proportionately across the remaining elements in series with it. For example, if a unit has 10 series groups of elements inside each unit and there are 5 units in a series string, then there will be 50 elements in series phase-to-neutral. If one element fails then the voltage on the remaining elements in series will increase by 1/50th of the total phase-to-neutral voltage. The unbalance protection would usually be set to trip when the overvoltage has exceeded the 110% capability of the elements.

When designing a fuseless bank and choosing the rated voltage and kvar of the units, keep in mind the continuous current capability and the transient current capability of the units. With the series string design of a fuseless bank, there are typically fewer parallel paths for these currents.

A typical fuseless shunt capacitor bank is shown in Fig. 10-114.

Harmonic filter shunt capacitor banks. Non-linear devices or loads, such as rectifiers and arc furnaces, can create harmonic distortion on the power system. Excessive harmonic voltages or currents can create problems for equipment and the electrical system in general. One common way to eliminate these harmonics is to install a passive harmonic filter near the device or load to shunt some of the harmonic current and thereby reduce the harmonic current flowing into the system. A schematic for a typical harmonic filter capacitor bank is shown in Fig. 10-115. The need for harmonic filters can be on low, medium, or high voltage electrical systems.

For low-voltage systems, the dry-type, metalized film capacitors are typically used and an inductor is used to create an LC circuit approximately tuned to the frequency of the harmonic being generated.

For medium and high voltage systems, the all-film, oil-filled dielectric capacitors are commonly used and, depending on the system voltage, a dry-type or oil-filled reactor is used to create an LC circuit approximately tuned to the frequency of the harmonic being generated. The capacitor portion for medium- and high-voltage harmonic filter banks can be externally fused, internally fused, or fuseless.

Harmonic filter capacitor banks may look similar to conventional shunt capacitor banks, but there are different criteria to consider when designing a harmonic filter bank. In general, the following should be considered when designing a harmonic filter bank:

1. Reactive power (kvar) requirements
2. Harmonic limitations
3. Normal system conditions, including ambient harmonics
4. Normal harmonic filter conditions
5. Contingency system conditions, including ambient harmonics
6. Contingency harmonic filter conditions

For detailed guidance in designing low, medium, and high voltage harmonic filter capacitor banks refer to IEEE 1531[3].

A typical harmonic filter shunt capacitor bank is shown in Fig. 10-116.

Protective Relaying for Shunt Capacitor Banks. Protection of shunt capacitor banks includes bank and system protection schemes.

Bank protection schemes usually involve some type of unbalance protection to trip the bank for overvoltage due to unit (element) failures, overcurrent protection for faults in the bank other than in units, and phase overcurrent or negative sequence protection for rack-to-rack flashovers in a phase-over-phase bank design.

System protection schemes can involve phase voltage relays for system overvoltage conditions, bus overcurrent, ground-overcurrent, or differential protection for faults in the capacitor installation

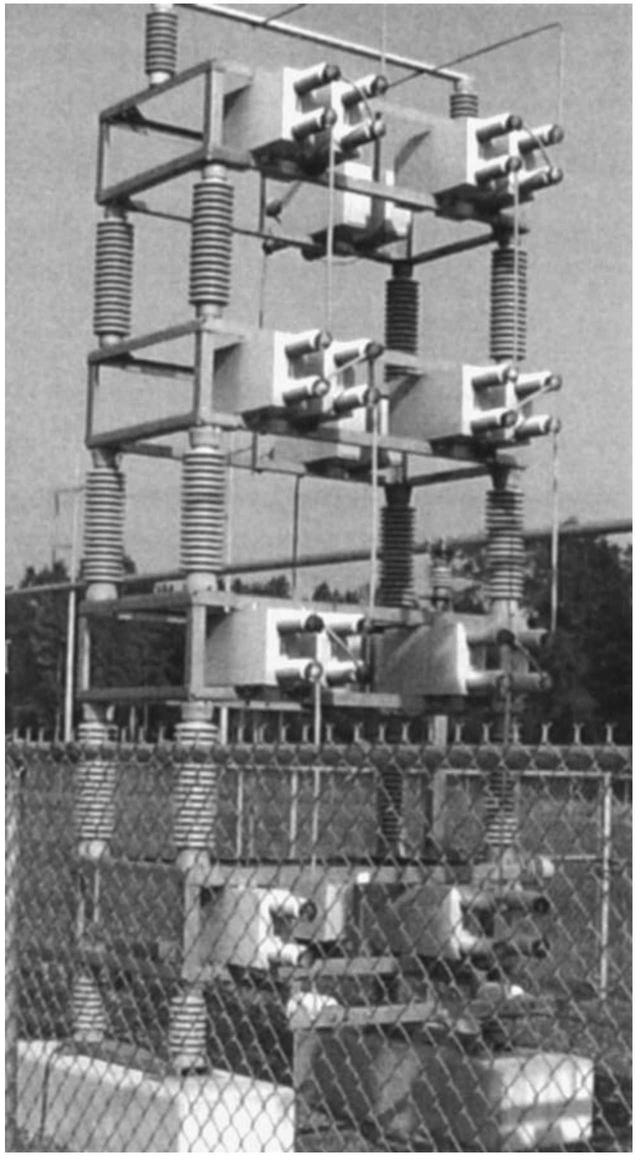


FIGURE 10-114 Example of a fuseless shunt capacitor bank.

or major capacitor bank failure, harmonic overload protection for excessive harmonic currents in the system, undervoltage relays for system outages, and breaker failure relays.

The following is a list of basic considerations for protection of shunt capacitor banks:

1. Type of shunt capacitor bank (externally fused, internally fused, or fuseless)
2. Bank connection (grounded wye, ungrounded wye, double wye, etc.)

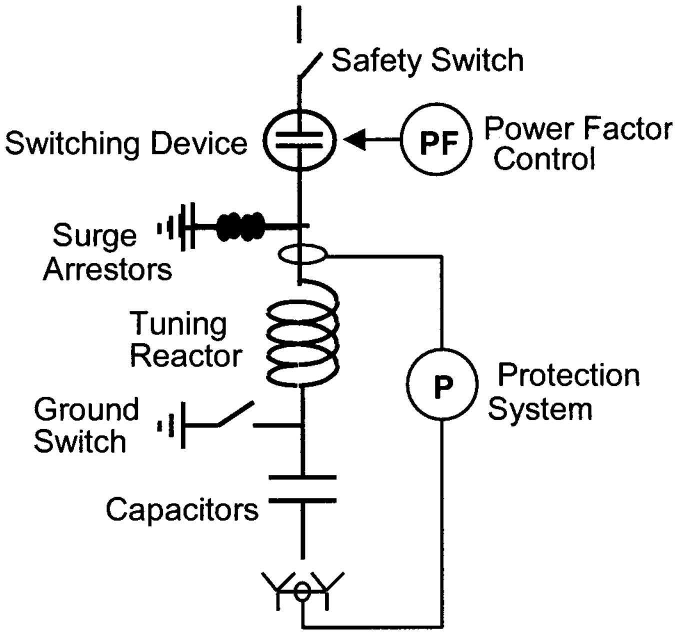


FIGURE 10-115 Typical schematic of a harmonic filter shunt capacitor bank.

FIGURE 10-116 Typical high-voltage harmonic filter shunt capacitor bank.

3. Capacitor bank design and arrangement of units
4. Normal and contingency system conditions
5. Point of connection to the power system
6. Capacitor unit capabilities
7. Overvoltage on remaining capacitor units
8. Installation grounding for grounded banks (single-point or peninsula grounding)

For detailed guidance on the protection of conventional and harmonic filter shunt capacitor banks refer to IEEE C37.99[6].

10.5.4 Series Capacitor Banks

Fixed (Conventional) Series Capacitors. A fixed series capacitor bank is an assembly of different components inserted in series with a power line to reduce the apparent reactance, increase the power transfer capability and improve system stability. The major components of a fixed series capacitor bank include capacitors, varistors, bypass gaps, bypass switches, discharge current limiting reactors, insulated structures, and protection and control systems. The inserted reactance is primarily established by the reactance of the capacitors. A schematic for a typical fixed series capacitor bank is shown in Fig. 10-117. The capacitors can be of the externally fused, internally fused, or fuseless design. Series capacitor banks are more commonly applied on higher voltage transmission systems, but they have been applied on distribution systems. See Fig. 10-118 for a typical high-voltage series capacitor bank.

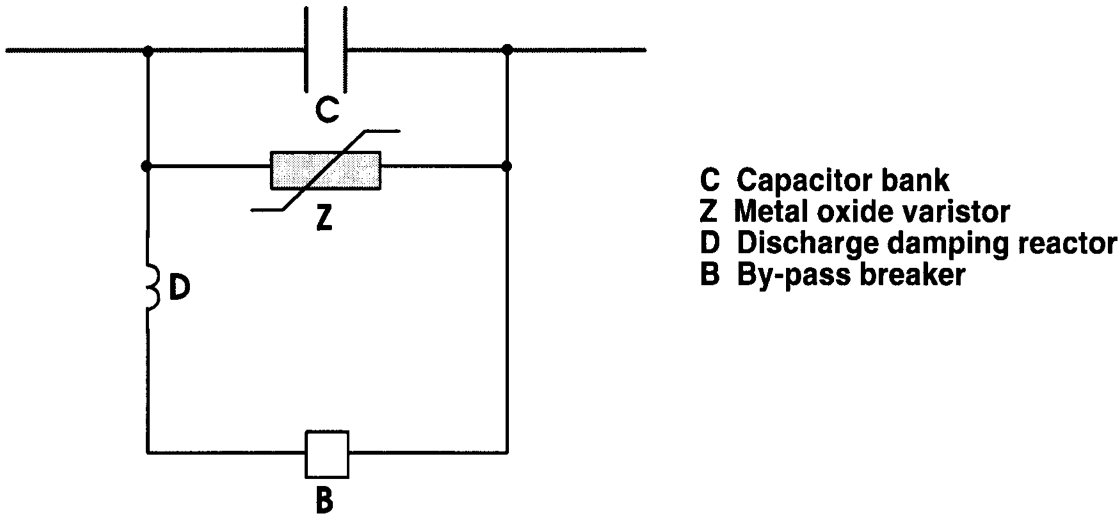


FIGURE 10-117 Schematic of fixed conventional series capacitor bank.

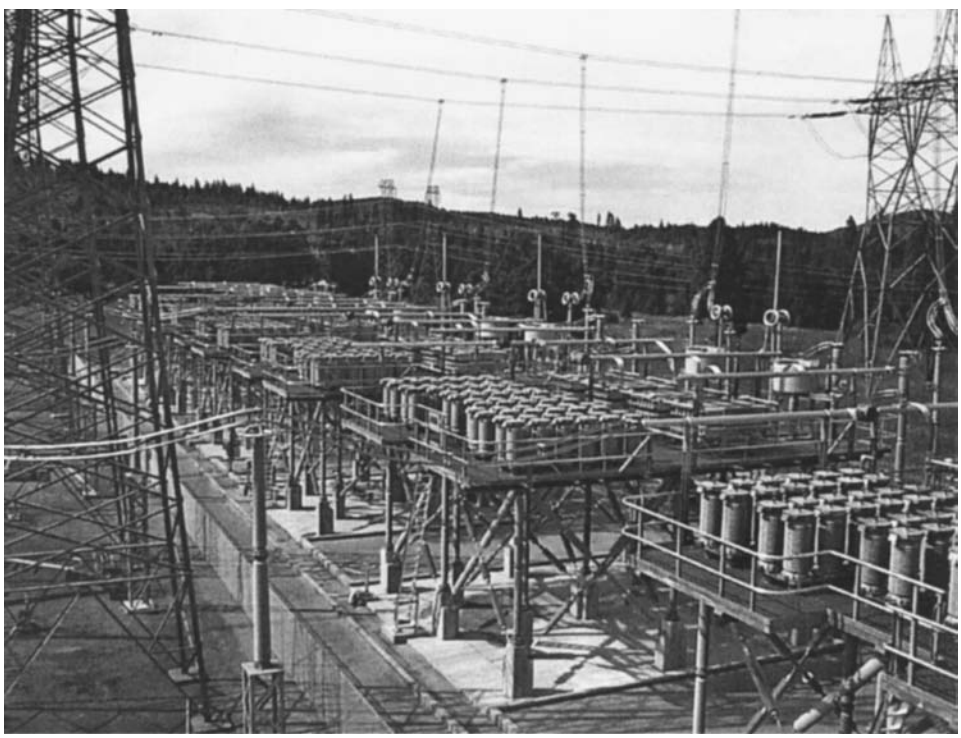


FIGURE 10-118 Typical fixed series transmission series capacitor bank.

For detailed information on the design, application, and protection of fixed transmission series capacitor banks, refer to IEEE 824[2,8]. There currently is also an IEEE working group preparing a guide for the specification of fixed transmission series capacitor banks.

For information on distribution series capacitors the user can refer to an IEEE Transactions paper written by the IEEE Series Capacitor Working Group[9].

General information about the protection of series capacitor banks can be found in the IEEE Special Publication TP-126-0[10]. There currently is also an IEEE working group preparing a guide for protection of transmission line series capacitors. When published, the document will be designated IEEE C37.116, *Guide for Protective Relay Application of Transmission-Line Series Capacitors*.

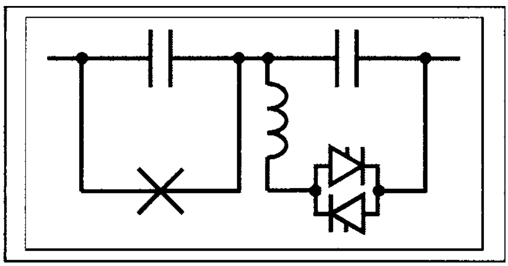


FIGURE 10-119 Typical schematic of a thyristor-controlled series capacitor bank.

Thyristor-Controlled Series Capacitors (TCSC). A thyristor-controlled series capacitor (TCSC) is a series capacitor bank paralleled with a thyristor-controlled reactor. In the controlled operation mode, periodic gating results in thyristor currents at a frequency higher than the power frequency which add with the line current. This boosts the capacitor voltage beyond the level that would be obtained by the line current alone. Also, controlled operation provides dynamic var compensation to control line current or through power. A schematic for a TCSC bank is shown in Fig. 10-119. The ability to control TCSC reactance rapidly provides a degree of transient stability and subsynchronous resonance mitigation, which is an advantage a TCSC has over a fixed series capacitor bank. A typical TCSC is shown in Fig. 10-120.

For detailed information on thyristor-controlled series capacitors, refer to IEEE 1534 [11].

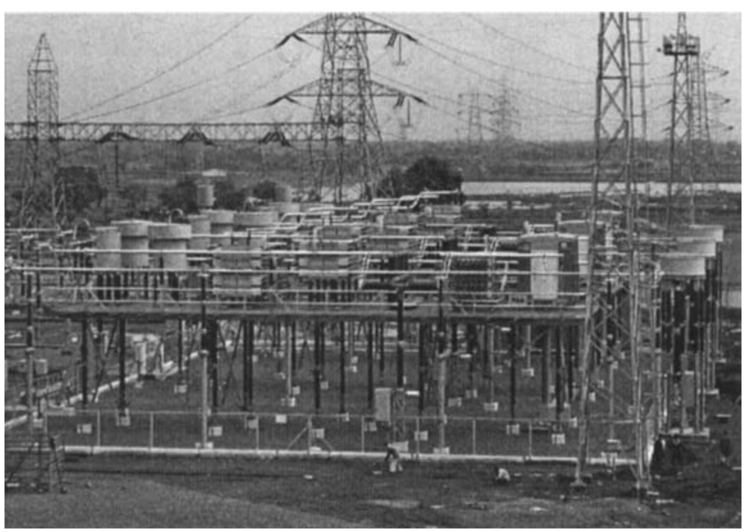


FIGURE 10-120 Typical thyristor-controlled series capacitor bank.

10.5.5 Capacitor Switching Equipment

This chapter will not go into the details of specifying equipment used to switch and protect capacitor banks. But a bibliography of documents providing rating, application, and specification information on switching equipment for shunt and series capacitors is provided at the end of this chapter for the user's reference.

REFERENCES

1. IEEE 18, *Standard for Shunt Power Capacitors.*
2. IEEE 824, *Standard for Series Capacitor Banks in Power Systems.*
3. IEEE 1531, *Guide for Application and Specification of Harmonic Filters.*
4. IEEE 1036, *Guide for the Application of Shunt Power Capacitors.*
5. IEEE C37.48, *Guide for the Application, Operation, and Maintenance of High-Voltage Fuses, Distribution Enclosed Single-Pole Air Switches, Fuse Disconnecting Switches, and Accessories* .
6. IEEE C37.99, *Guide for the Protection of Shunt Capacitor Banks.*
7. IEEE C37.48.1, *Guide for the Operation, Classification, Application, and Coordination of Current-Limiting Fuses with Rated Voltages 1-38 kV.*
8. P. M. Anderson and R. G. Farmer, *Series Compensation of Power Systems*, PBLSH! Inc., 1996, Internet-http://www.pblsh.com/CDsBooks.
9. "Considerations for the application of series capacitors to radial power distribution circuits," IEEE Transactions on Power Delivery, vol. 16, Issue 2, April 2001, Pages:306–318.
10. IEEE Special Publication TP-126–0, *Series Capacitor Bank Protection.*
11. IEEE 1534, *Recommended Practice for Specifying Thyristor Controlled Series Capacitors.*

BIBLIOGRAPHY

Capacitor Literature

Power Capacitor Bibliography, IEEE/PES Capacitor Subcommittee web page, http://grouper.ieee.org/ groups/td/cap/

Power Capacitor for Harmonic Filter Bibliography, IEEE/PES Capacitor Subcommittee web page, http://grouper.ieee.org/groups/td/cap/

Series Power Capacitor Bibliography, IEEE/PES Capacitor Subcommittee web page, http://grouper.ieee.org/ groups/td/cap/

Benitez, J., *Application of capacitors for power factor correction of industrial electrical distribution systems,* Industry Applications Society 39th Annual Petroleum and Chemical Industry Conference, Record of Conference Papers., 28–30 Sept. 1992, Page(s):77−86.

Bonner, J.A., Hurst, W.M., Rocamora, R.G., Dudley, R.F., Sharp, M.R., Twiss, J.A., *Selecting ratings for capacitors and reactors in applications involving multiple single-tuned filters,* IEEE Transactions on Power Delivery, vol. 10, Issue 1, Jan. 1995, Page(s):547−555.

Dugan, R.C., Kennedy, B.W., *Predicting harmonic problems resulting from customer capacitor additions for demand-side management,* IEEE Transactions on Power Systems, vol. 10, Issue 4, Nov. 1995, Page(s):1765–1771.

Hammond, P.W., *A harmonic filter installation to reduce voltage distortion from static power converters,* IEEE Transactions on Industry Applications, vol. 24, Issue 1, Part 1, Jan.-Feb. 1988, Page(s):53–58.

McGranaghan, M.F., Mueller, D.R., *Designing harmonic filters for adjustable-speed drives to comply with IEEE-519 harmonic limits,* IEEE Transactions on Industry Applications, vol. 35, Issue 2, March-April 1999, Page(s):312–318.

Mendis, S.R., Bishop, M.T., McCall, J.C., Hurst, W.M., *Overcurrent protection of capacitors applied on industrial distribution systems,* IEEE Transactions on Industry Applications, vol. 29, Issue 3, May-June 1993, Page(s):541–547.

Miske, S.A.(Working Group paper of the IEEE Capacitor Subcommittee), *Considerations for the application of series capacitors to radial power distribution circuits,* IEEE Transactions on Power Delivery, vol. 16, Issue 2, April 2001, Page(s):306–318.

Peggs, J.F., Powell, P.W., Grebe, T.E., *Innovations for protection and control of high voltage capacitor banks on the Virginia power system,* Proceedings of the 1994 IEEE Power Engineering Society Transmission and Distribution Conference, 10–15 April 1994, Page(s):284–290.

Peeran, S.M., Cascadden, C.W.P., *Application, design, and specification of harmonic filters for variable frequency drives,* IEEE Transactions on Industry Applications, vol. 31, Issue 4, July-Aug. 1995, Page(s): 841–847.

Sevigny, R., Menard, S., Rajotte, C., McVey, M., *Capacitor measurement in the substation environment: a new approach,* 2000 IEEE ESMO-2000 IEEE 9th International Conference on Transmission and Distribution Construction, Operation and Live-Line Maintenance Proceedings, 8–12 Oct. 2000, Page(s):299–305.

Thiel, P.H., Harder, J.E., Taylor, G.E., *Fuseless capacitor banks,* IEEE Transactions on Power Delivery, vol. 7, Issue 2, April 1992, Page(s):1009–1015.

Standards for Equipment used to Switch Power Capacitors

IEEE 824, *Standard for Series Capacitor Banks in Power Systems.*

IEEE 1247, *Standard for Interrupter Switches for Alternating Current, Rated above 1000 volts.*

IEEE C37.04, *Standard Rating Structure for AC High-Voltage Circuit Breakers.*

IEEE C37.012, *Application Guide for Capacitance Current Switching for AC High-Voltage Circuit Breakers.*

IEEE C37.016, *Standard for Circuit Switchers.*

IEEE C37.66, *Standard Requirements for Capacitor Switches for AC Systems, 1-38 kV.*

IEC 62271-109, *Alternating current series capacitor by-pass switches.*

10.6 FUSES AND SWITCHES

BY JON HILGENKAMP

10.6.1 Fuses

General Classifications. Fuses can be classified into three categories. The first is low-voltage fuses operating up through 600 V ac. Most devices are tested and approved by Underwriters' Laboratories, Inc., and are marketed in a wide variety of characteristics and physical configurations. The second classification is medium-voltage through 69 kV, and the third is high-voltage fuses through 500 kV. These fuses are used by electric utilities to protect transmission-distribution-class equipment and by large industrial complexes which have their own electrical distribution systems.

Low-Voltage Fuses. Low-voltage fuse standards are established by ANSI, NEMA, or Underwriters' Laboratories (see Table 10-13). Their characteristics include a voltage class, an ampere rating, and an interrupting rating and, for some classes of fuses, a current-limiting rating. Cartridge fuses are classified in the following voltage classes: not over 250 V ac, not over 300 V ac, and not over 600 V ac. Fuses should not be used for dc applications unless recommended by the manufacturer. Most low-voltage fuses must be used in accordance with the National Electrical Code. Exceptions include those

TABLE 10-13 Class of Low-Voltage Fuses and Applicable Standards

Fuse class	Standard
Plug fuses	UL 198F
Radio and appliances	UL 1417
Mine duty	UL 198M
DC fuses	UL 198L
H	UL 198B and ANSI C97.1
G, J, L	UL 198C and ANSI C97.1
K	UL 198D and ANSI C97.1
R	UL 198E
T	UL 198H

used in ships, railways, aircraft, and automotive vehicles other than mobile homes and recreational vehicles.

The standard lines of low-voltage fuses are available in several steps of ampere capacity, each of which is a different physical size (see Table 10-14).

TABLE 10-14 Typical Dimension Grouping of Fuses

Class	Volts	Amperes
G	300	0–15
G	300	16–20
G	300	21–30
G	300	31–60
H, K	250, 600	0–30
H, K	250, 600	31–60
H, K	250, 600	61–100
H, K	250, 600	101–200
H, K	250, 600	201–400
H, K	250, 600	401–600
L	600	601–800
L	600	801–1200
L	600	1201–1600
L	600	1601–2000
L	600	2001–2500
L	600	2501–3000
L	600	3001–4000
L	600	4001–5000
L	600	5001–6000

A minimum of 10,000-A interrupting capacity is typical in low-voltage fuses, but some sizes and types are capable of interrupting up to 200,000 A ac or 100,000 A dc. The interrupting rating is the highest rms symmetric alternating current which the fuse can interrupt at rated voltage.

Low-voltage current-limiting fuses are designed so that non-current-limiting fuses cannot be inserted into the fuse holder as a direct replacement. Thus, Class K fuses which are interchangeable with Class H fuses are not permitted the "current-limiting" label.

A low-voltage current-limiting fuse successfully and safely interrupts all available currents within its specified interrupting rating and, within its current-limiting range, limits the clearing time at rated voltage to an interval equal to or less than the first major current loop. These fuses also limit peak let-through current to a value less than the normal peak current that would be possible without current-limiting availability. The current-limiting characteristics of current-limiting fuses are expressed by two electrical measurements: (1) maximum peak let-through current, the maximum instantaneous value of current passed by the fuse during time of operation; and (2) maximum clearing I^2t (amperes-squared-seconds), an expression of the energy available as a result of current flow during the clearing time of operation.

Limiters are time-delay fusible connectors designed to be installed in low-voltage network-mains cables at street-junction points. They are rated in cable sizes and have time-current characteristics to (1) allow the cable fault to burn itself clear if it does so promptly, without blowing the limiter; (2) blow before the cable insulation away from the fault is damaged and prevent the failure from spreading beyond the junction; and (3) obtain adequate selectivity so that, when installed in a network, only those limiters connected to the faulted cable blow (see Fig. 10-121 for limiter characteristics).

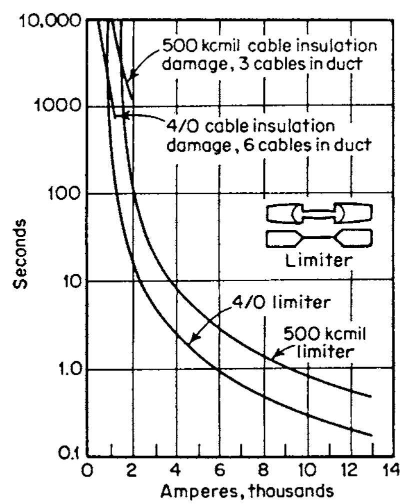


FIGURE 10-121 Time-current characteristics of low-voltage limiters for No. 4/0 and 500 Mcmil network cables.

The usual applications of limiters are at the street-corner junctions of low-voltage network mains, multiple-transformer secondary cables, multiple-service cables, and intervault ties of building interior networks.

In recent years, combinations of low-voltage circuit breakers and low-voltage current-limiting fuses have been devised to provide improved short-circuit protection, especially on motor branch circuits where contactors have low interrupting ratings.

Medium- and High-Voltage Fuses. A medium-voltage fuse is defined as any fuse (above 600 V and less than 69 kV) or fuse device used to isolate an electric short circuit from an electrical distribution system. A *high-voltage fuse* is defined as any fuse rated above 69 kV and less than 230 kV used for this same purpose. Specific classes of fuses or fused devices are

Enclosed cutouts and fuses
Open cutouts and fuses
Open-link cutouts and fuses
Current-limiting fuses
Power fuses
Oil-immersed protective links

See Table 10-15.

TABLE 10-15 Classes of High-Voltage Fuses or Fused Devices and Applicable Standards

Class device	Standard
Distribution cutouts and fuse links	ANSI C37.42
Distribution oil cutouts and fuse links	ANSI C37.44
Power fuses	ANSI C37.46
Current-limiting fuses	ANSI C37.47

These fuses are used to protect potential transformers, distribution or power transformers, and lateral taps from main distribution feeder circuits. They are often used as sectionalizing devices on main feeder circuits. The ampacity and interrupting rating of these devices range up to 720 A, 500 mVA at 14,400 V, and 250 A, 2000 mVA at 138,000 V. Fuses are generally used in electrical series with other fuses or circuit-protective devices. Care must be taken in coordinating the time-current characteristics for proper isolation of the electric circuit during fault and overload conditions.

Distribution fuse links for use with expulsion cutouts are available with many different time-current characteristics. Figure 10-122 shows the minimum melting time-current characteristics for ANSI Type K fuse links.

In an effort to standardize fuse-link characteristics, ANSI has adopted time-current characteristics for two basic fuse-link types: Type K (fast) and Type T (slow). These fuse links are designed so as to have the same time-current characteristics regardless of manufacturer. A wide variety of non-standardized fuse-link characteristics are also available. Fuse-link characteristics are usually based on tests at an ambient temperature of 25°C and no initial load. For characteristics at other ambients or for preloading variations, consult the individual manufacturer.

The characteristics of various medium- and high-voltage protective devices used on distribution power systems are

1. Expulsion cutouts
   a. *Open type:* 20-kA asymmetric maximum interrupting current (IC), 5 to 35 kV, violent in operation at high faults, low cost, both initial and refusing. Maximum continuous-current rating of 200 A; can be fitted with a solid blade for conversion to a disconnect switch with a rating of 300 A. Figure 10-123 shows a typical open-type cutout.
   b. *Enclosed type:* 8- to 10-kA asymmetric maximum interrupting current, used primarily where safety codes dictate use. Enclosed cutouts are no longer manufactured and are being replaced with open-type cutouts.
   c. *Open link:* 1200-A maximum interrupting current, 50-A maximum continuous current, applied on rural lines and/or small transformers.

2. *Oil cutouts:* Considerable application in the past, especially in underground vaults; however, low interrupting current now poses serious underrating problems.

3. *Liquid fuses:* Nonviolent, low interrupting current (8 to 10 kA maximum).

4. *Power fuses:* Reduced arc energy, somewhat less violent than cutouts on high faults, rated to 20-kA interrupting current; both initial purchase and replacement expensive. Figure 10-124*a* shows an indoor-type power fuse.

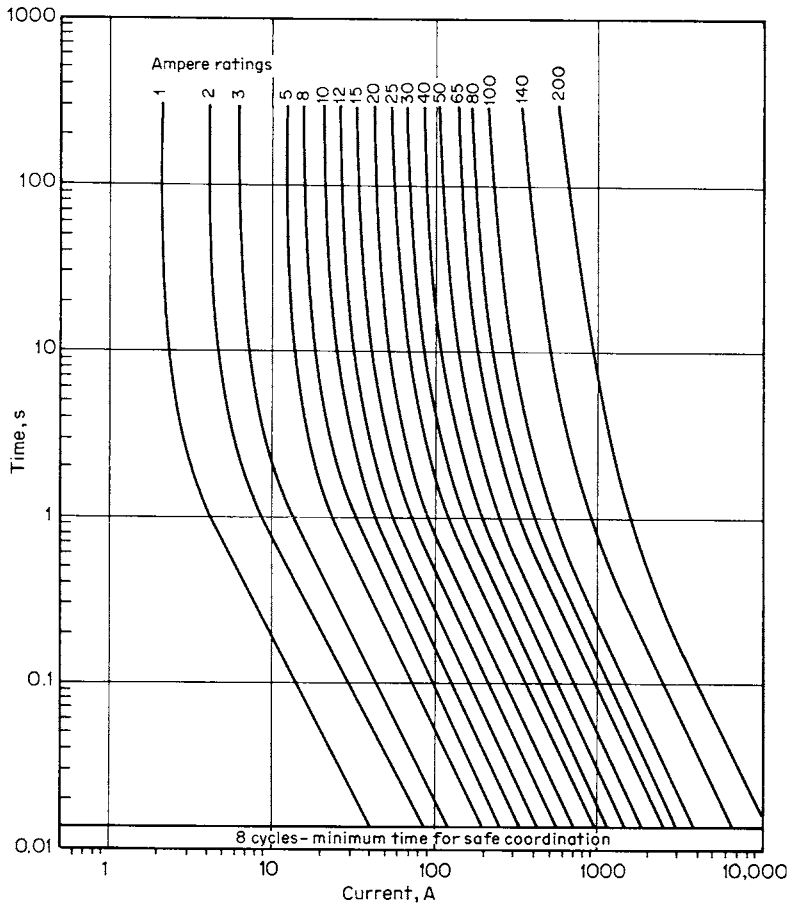


FIGURE 10-122 Minimum melting time-current characteristics of a NEMA Type K fuse. *(A. B. Chance Company.)*

5. *Under-oil protective link:* 4000-A asymmetric maximum interrupting current, violent in operation, low cost, contaminates insulating oil.
6. *General-purpose current-limiting fuses:* Nonviolent, current-limiting, high interrupting current (50 kA), requires coordination study, generated peak arc voltage, not affected by system transient recovery voltage, both initial purchase and replacement expensive.
7. *High-range backup current-limiting fuses:* Current-limiting, high interrupting current (50 kA), requires a low-current interrupting device in series, operates only at high currents, does not affect existing system coordination, low refusing cost on majority of outages because of only blowing expulsion link, not affected by system transient recovery voltage.
8. *Vacuum fuses:* Function with no external arcing or violence. They are nonrenewable and the associated cost is high. They have not found widespread application.
9. *SF_6 fuses:* Function with no external arcing or violence. They utilize rotating-arc technology and have a 12.5-kA interrupting current rating.

All expulsion-principle fuses depend on arc-quenching material—bone fiber, liquid solutions, or boric acid powder—to develop water vapor and/or other gases to cool the arc from the melted fuse

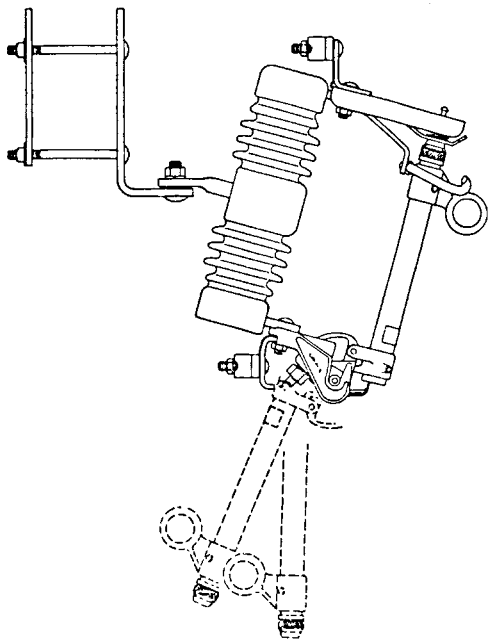


FIGURE 10-123 Open-type distribution fuse cutout. *(A. B. Chance Company.)*

link. These fuses have no energy-limiting ability and require a natural current-zero crossing to successfully interrupt a short-circuit current. Figure 10-124*b* shows a cross section of a power-type fuse.

Current-Limiting Fuses. High-voltage current-limiting fuses for use on distribution systems have two distinct classes: (1) *general-purpose current-limiting fuses,* devices which will successfully interrupt currents which will melt the fusible element in 1 h—these are sometimes referred to as full-range fuses, but this is not an industry-accepted designation; and (2) *backup current- limiting fuses,* devices which have a definite minimum interrupting rating as specified by the manufacturer. These devices require other protective devices in electrical series to protect the backup current-limiting fuse from currents below its minimum interrupting rating. Recently, a new type of backup current-limiting fuse has entered the market. This device, which is referred to as a *fuse limiter,* combines the series fuse and the current-limiting fuse into one convenient package that handles just like a cutout. The fuse limiter is available at 15 and 25 kV, and in current ratings through 20 A. It is specifically designed for protecting overhead distribution transformers (see Fig. 10-125).

Three important parameters should be known about high-voltage current-limiting fuses:

1. *Continuous current rating:* The maximum current that the fuse is designed to carry continuously.
2. *Peak arc voltage:* Maximum voltage generated by the current-limiting fuse. If wire-wound, the voltage value is a function of fault current. If it is a ribbon-element fuse, the voltage is a function of applied voltage across the fuse.
3. *I^2t clearing:* Maximum allowed by the current-limiting fuse. This measures the energy-limiting effect of the fuse.

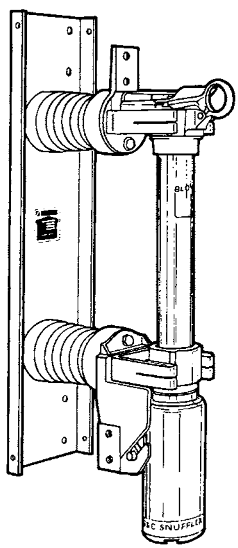
Care in application of current-limiting fuses according to voltage rating must be maintained. In general, these fuses should not be applied to circuits with a voltage less than 50% of the fuse-voltage rating to avoid excessive peak arc voltages.

It is equally important that fuses not be exposed to system recovery voltages in excess of their rating.

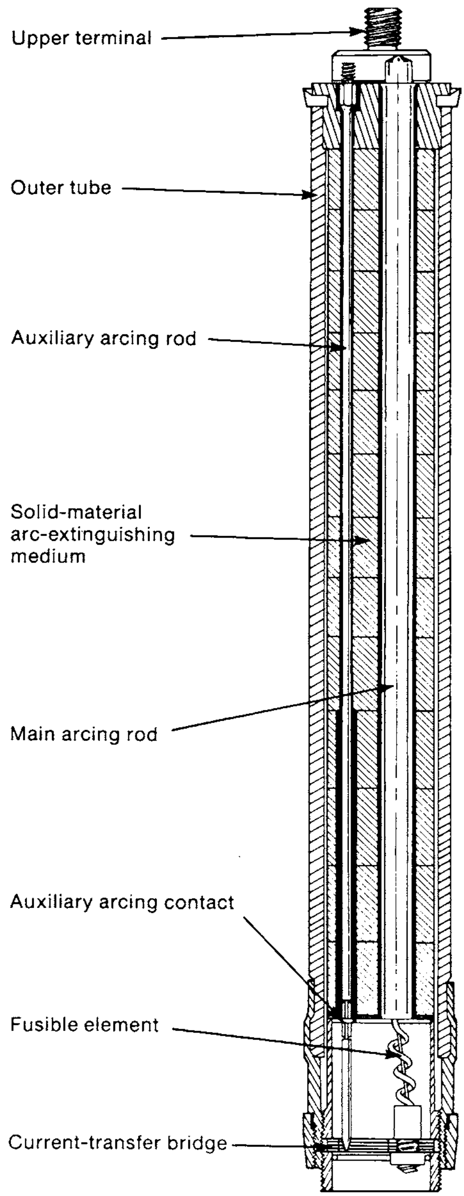
Fuses in Enclosures. High-voltage fuses may be mounted in enclosures for several applications: industrial service entrance switchgear, pad-mounted switchgear or transformers for underground circuits, or in enclosures for subsurface applications. Most fuses will require special adaptation. Power fuses are fitted with a muffler to reduce the intensity of the exhaust gases when used in enclosures. Current-limiting fuses are supplied with special seals to prevent the ingress of fluid when applied under oil such as in transformers. Fuse cutouts are not recommended for use in enclosures or vaults.

Derating of fuses in enclosures may have to be considered because of restricted heat transfer. Consult the manufacturer.

Electronically Controlled Protective Devices. Electronically controlled protective devices offer greater flexibility and accuracy through the use of state-of-the-art electronics. An electronic power



(a)



(b)

FIGURE 10-124 (*a*) Indoor-type power fuse; (*b*) cross-sectional view of power-fuse refill unit. *(S&C Electric Company.)*

fuse and an electronic sectionalizer for use at distribution voltage levels are two examples of this technology.

The electronic power fuse utilizes an electronic control module to provide current sensing, time-current characteristics, and control power for the fuse. High-speed interruption of fault currents to 40,000 A is provided by an interrupting module. The devices are completely self-contained (nonventing) and

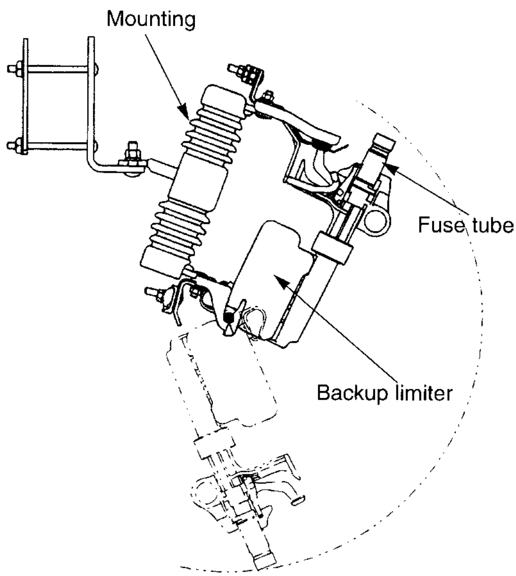


FIGURE 10-125 Fuse limiter. *(S&C Electric Company.)*

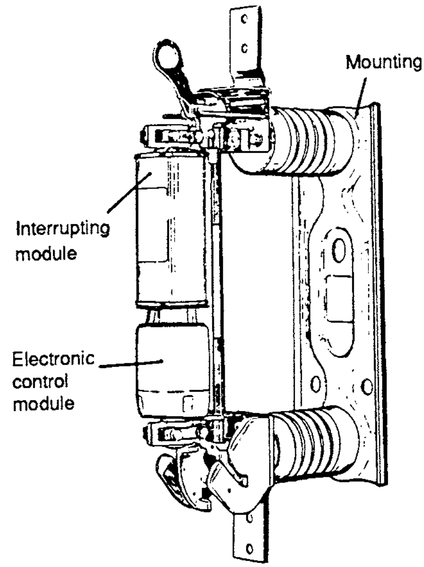


FIGURE 10-126 Electronic power fuse. *(S&C Electric Company.)*

require no external power source. These power fuses are used in metal-enclosed switchgear, pad-mounted switchgear, and metal-enclosed fuse gear. They are available to 600 A continuous current and 4.16 through 25 kV. Three families of time-current characteristics are available. Figure 10-126 shows an electronic power fuse.

The electronic sectionalizer consists of an electronic module which fits into the mounting of a standard open-type cutout. The sectionalizer counts the number of fault current pulses allowed by an upstream recloser and operates to an open position during the recloser's open time. The sectionalizer has no time-current characteristics. It is generally used to replace a fuse at the distribution lateral where fuse coordination is difficult or impossible. Sectionalizers have no fault-interrupting capability and must be used in conjunction with an upstream recloser. Sectionalizers are available in continuous current ratings to 200 A; count settings of 1, 2, or 3; and distribution voltage ratings from 15 to 38 kV. The electronic modules require no external power source. An electronic sectionalizer is shown in Fig. 10-127.

10.6.2 Switches

Disconnecting switches are used primarily for isolation of equipment such as buses or other live apparatus. They are used for sectionalizing electric circuits such as buses or lateral circuits or even portions of main feeders for special purposes such as testing and maintenance. Standards pertaining to disconnect switches are listed in Table 10-16. Generally, these devices are not rated to break load current except when equipped with specific auxiliary devices. Interruption of load currents, magnetizing currents, or capacitive currents without other aiding devices may not be successful. While not recommended, disconnect switches without loadbreak capability have been used to interrupt limited values of load, charging, and magnetizing current. Their performance depends not only on the magnitude of the current but also on the speed of operation, wind direction, velocity, and clearance to other energized or grounded components. The principal concern is that the arc which is created may establish a high-current fault. Arc reach may approach tens of feet or more depending upon system voltage and the current being interrupted. However, these switches must be designed to carry

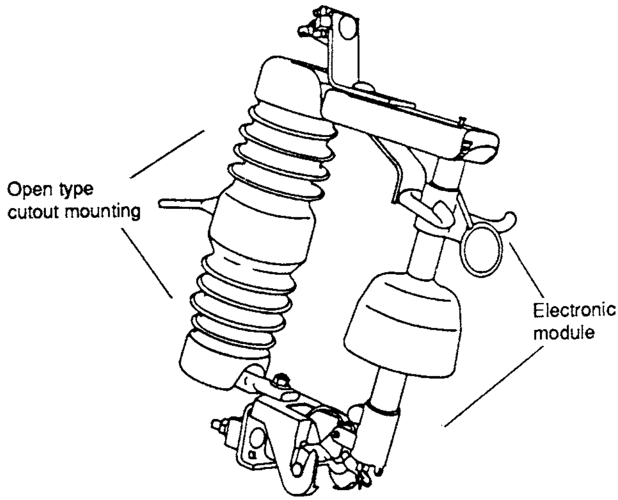


FIGURE 10-127 Electronic sectionalizer. *(S&C Electric Company.)*

expected load currents and remain closed for momentary current flow such as fault currents. Fault currents in excess of a specific rating may cause the switch to be blown open by the magnetic forces due to the short-circuit current.

There are three classes of disconnect switches:

1. Station
2. Transmission
3. Distribution

Switches can be further categorized as gang- or hookstick-operated and loadbreak or nonloadbreak types.

The basic insulation level (BIL) of station-class equipment is normally higher than for transmission or distribution equipment. Station equipment ranges from 2.4 to 765 kV at present. Disconnect switches rated up through 3000 A are available. Manual or automatic switching can be provided.

The design of disconnecting switches demands considerable attention to the contact surfaces. Consideration must be given to the rigors of extreme environments.

High-pressure contacts are generally the form used to provide the current transfer. Current densities of 100,000 A/in^2 are not uncommon when using silver for contact points. Contact pressures as high as 500,000 lb/in^2 ensure that good cleaning action is achieved and keeps the current transfer points free from contamination.

Transmission disconnecting switches are generally used as load-management tools. Increasing needs for transmission lines and decreasing availability of right-of-way makes automatic switching of transmission load desirable.

Load management is achieved during "dead time" by switching the proper disconnects automatically through sensing loss of voltage. These systems are available through 161 kV, 1200 A.

The objective of load management is to minimize outage time and allow for more efficient utilization of substation capacity at the distribution level. There is a growing interest in automating distribution class switches to achieve load-management objectives.

TABLE 10-16 Standards Related to Disconnect Switches

Ratings and Application Guide	ANSI C37.32
Rated Control Voltages	ANSI C37.33
Test Code	ANSI C37.34
Operation and Maintenance	ANSI C37.35
Loading Guide	ANSI C37.37

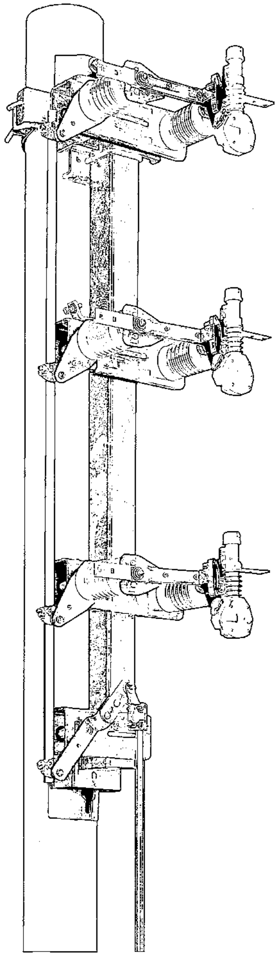


FIGURE 10-128 Gang-operated, phase-over, distribution loadbreak switch. *(S&C Electric Company.)*

Distribution disconnecting switches are becoming the method of providing for both single- and 3-phase sectionalizing. As distribution voltage grows to higher levels, the utility must provide more sectionalizing or switching capability or suffer large and longer outages during faults. Hence, at 34.5 kV one may find some switching capability every 2000 to 3000 ft of overhead conductor. Figure 10-128 shows a typical gang-operated distribution loadbreak switch.

Single-phase disconnect switching with load-interrupter capability can be applied where ferroresonance is not a problem. This type of switching is found on single-phase circuits. Single-phase switching of a heavily loaded 3-phase circuit is not desirable.

Gang-operated switches with loadbreak capability interrupt these loads without concern for ferroresonance problems. In application, these 3-phase switches can be mounted in either horizontal or vertical configurations. To ensure proper operation, the mounting should be as rigid as possible. Care must be exercised in proper alignment of blades and clips. Attention to these matters allows proper operation without overstressing porcelain insulators or distortion of contact surfaces and operational handles.

Load-Interrupter Devices. Load-interrupter devices, when combined with disconnecting switches, provide the desired capability of switching load currents without circuit breakers. Generally, these interrupters are auxiliary devices and are not continuous-duty in terms of carrying load. This load interruption can be achieved by

1. Insertion of a resistor in the circuit following opening of the main switch contact and interruption of the current drawn between arcing horns in the air.
2. Use of a blast of air or other gas to effectively lengthen the arc resulting from the main contacts opening.
3. Use of an interrupter paralleling the main contacts just prior to opening and interrupting in this auxiliary chamber after the main contacts open. This is typically accomplished with an expulsion-type device or vacuum switch.

Figure 10-129 shows an expulsion-type load interrupter used on distribution disconnecting switches to assist in interrupting load current.

Switches for Underground Circuits. The continuing trend toward underground distribution circuits increases the need for pad-mounted switchgear. These are available in both live- and dead-front configurations, with the latter growing in popularity.

Live-front switchgear is typically air-insulated and utilizes in-air switches for loadbreak operation and power fuses for fault interruption. All components are directly accessible and operable.

Dead-front switchgear is typically air-, oil-, or gas-insulated. Air-insulated dead-front switchgear is similar to live-front switchgear except that components are isolated within grounded compartments and are not accessible when energized.

Oil-insulated dead-front switchgear typically uses in-oil switches or vacuum interrupters for loadbreak operation and current-limiting fuses or vacuum interrupters for fault interruption. These have the advantage of being more compact than air-insulated units. The disadvantage is that the oil insulation can become contaminated following in-oil arc interruption or from external contaminants. SF_6-insulated switchgear uses components similar to those used in oil-insulated switchgear. Contamination of SF_6 is less of a concern, although SF_6-insulated gear must be more carefully designed and constructed to ensure gas integrity.

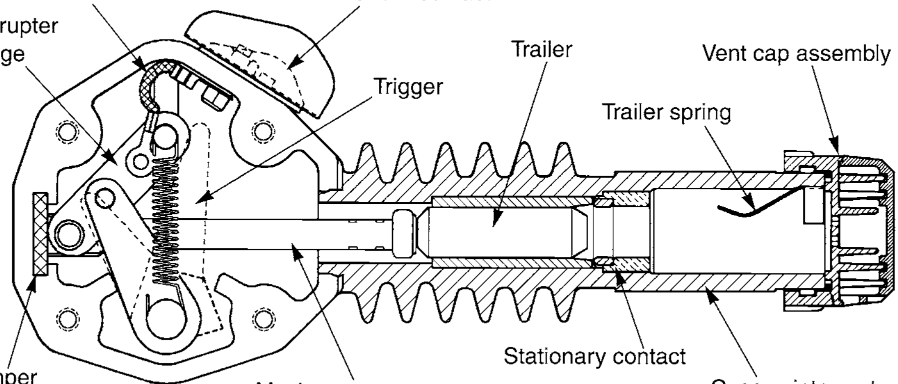


FIGURE 10-129 Load interrupter, expulsion type, used in distribution systems. *(S&C Electric Company.)*

Dead-front designs generally provide increased isolation from energized components, but at the expense of operating simplicity and visual confirmation. External connections are made by means of separable insulated connectors. These connectors generally must be removed to provide a visible break when working on cable or equipment. Some more recent dead-front designs allow visual confirmation of internal visible breaks on both switches and fault interrupters, and some provide integral grounding of cables. This eliminates the need to move elbows and provides isolation from energized components, while also offering the visual confirmation provided in live-front switchgear.

In addition, recent designs have become available either with provisions for motor operation or fully integrated with motor operation and controls for use with SCADA systems.

Oil-insulated and SF_6-insulated switchgear can also be submersible and thus used in subsurface applications where space is at a premium or aesthetics are critical, such as in metropolitan areas. Refer to ANSI C37.72 for standards governing dead-front pad-mounted switchgear, and ANSI/IEEE C37.71 for subsurface load-interrupting switches.

10.7 CIRCUIT SWITCHERS

BY JON HILGENKAMP

Circuit switchers are mechanical switching devices suitable for frequent operation; not necessarily capable of high-speed reclosing; capable of making, carrying, and breaking currents under normal circuit conditions; capable of making, and carrying for a specified time, currents under specified abnormal conditions; and capable of breaking currents under certain other specified abnormal circuit conditions. They may include an integral isolating device. Circuit switchers available today use SF_6 as an interrupting medium and may be equipped with a trip device connected to a relay to open the circuit switcher automatically under specified abnormal conditions, such as overcurrent or faults.

A circuit switcher, like a circuit breaker, must carry normal load currents within a specified temperature range to prevent damage to key components such as contacts, linkage, terminals, and isolating device parts. Principal designating parameters of a circuit switcher are maximum operating voltage, BIL, rated load current, interrupting current, whether an isolator is required, whether a trip device is required, and whether manual or motorized operation is required.

A circuit switcher essentially combines the functions of a circuit breaker (without reclosing capability) and a disconnecting switch (by providing visible isolation, but not necessarily meeting the safety requirements of all users). A circuit switcher provides a cost-effective alternative means of transformer protection and switching, line and loop switching, capacitor or reactor switching, and load management, with protection in most instances.

Evolution of the circuit switcher concept provides a more in-depth understanding of its application versatility and its limitations.

10.7.1 History of Circuit-Switcher Development

After World War II, the drive to electrify the remaining rural and sparsely populated areas of the United States was renewed. Providing fully rated circuit breakers for switching loaded circuits was frequently beyond budget limitations. This created a need for new transmission and subtransmission voltage circuit-switching devices. One such device could be described as a load interrupter. It appeared in a wide variety of forms. Most were attachments to disconnect switches.

Initially, most of these devices used low-volume oil as an interrupting medium. Ablative gas-generating devices and later vacuum displaced oil. With rare exceptions, these devices had deficiencies.

In the mid-1950s, SF_6 was first employed as an interrupting medium. The application was an interrupter attachment for disconnect switches.

Whereas ablative devices and vacuum bottles are limited to approximately 30-kV recovery voltage per gap, this single-gap SF_6 device was readily applied on 138-kV systems for up to 600 A loadswitching. Most of these vacuum, ablative, and SF_6 devices were shunted into the circuit during the disconnect switch opening process. As the 1960s approached, the circuit switcher was born. It appeared as an in-line device. While the first version employed a number of ablative devices in series, it soon evolved into the use of SF_6 as a medium. Because of the unfavorable experience with the earlier devices, the general acceptance of the circuit switcher took much effort and considerable time. A typical installation is shown in Fig. 10-130.

Applications for circuit switchers have been primarily for transformer protection. The circuit switcher provides load-switching capability and mainly protection for faults that originate on the secondary side of the substation transformer. The zone of protection for circuit switchers in this

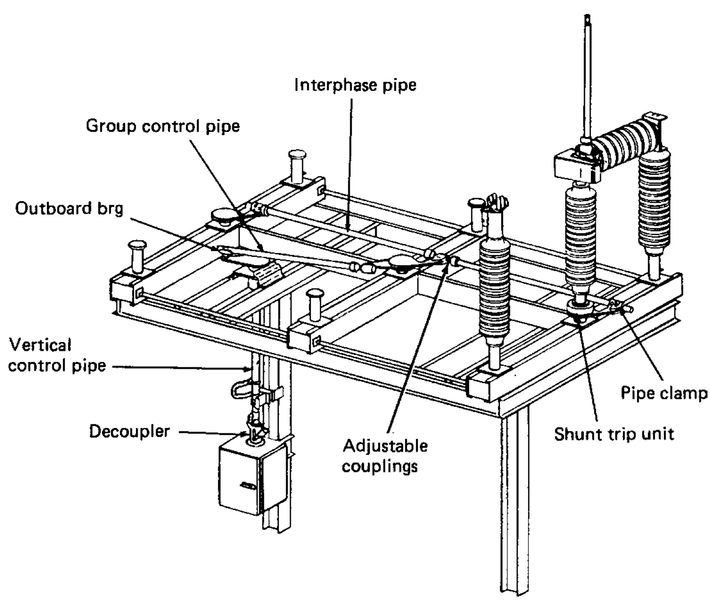


FIGURE 10-130 Schematic of typical 3-pole arrangement. Two poles have been deleted to clarify mechanical drive-train arrangement.

location is typically from the current transformers inside the transformer on the high-voltage bushings to the secondary feeder breakers. There is generally shorter strike distance on the secondary bus and more exposure to flashover from wildlife and other causes. Therefore, circuit switchers are specifically tested to interrupt the higher transient recovery voltages (TRVs) associated with faults initiated on the secondary of the transformer and cleared by the high-side protective device. For application where the available high-side short-circuit current exceeds the device's capability, blocking relays can be used. However, in most applications this is not necessary.

10.7.2 General Construction

Live-tank SF_6 gas puffer-type interrupters are utilized by most circuit switchers today. In the closed position, the contacts are surrounded by a flow guide and piston assembly which is ready to mechanically generate a "puff" of SF_6 to cool and deionize the arc that is established prior to circuit interruption. The moving cylinder attached to the contact assembly is driven by the main opening spring, causing the gas to be pressurized by the stationary piston. The stationary contact "follows" the moving contact as the piston assembly achieves the prepressurized gas condition. When the contacts (which are hollow tubes) part, an arc is established and the gas flow divides into two parts and flows down the stationary and moving contact tubes. The alternating nature of the arc current waveform results in two current zeros every cycle. As long as the arc is sufficiently "hot" or conductive through the SF_6 dielectric medium, the current will reestablish. At the first current zero where the SF_6 density is sufficient to stop the arc from reestablishing itself and to provide necessary dielectric strength, the arc is interrupted. This entire process from trip signal initiation to current interruption requires from 3 to 8 cycles or 50 to 133 ms in modern circuit switchers.

Figure 10-131 illustrates a typical "blade-disconnect model" circuit switcher with the interrupter and blade connected in series. For opening, the trip device, called a "shunt trip," receives a trip signal when the relay system detects an abnormal condition within the specified range or when the operator desires a high-speed circuit opening. By discharging its operating spring, the shunt trip rotates the insulator above it at high speed, thus tripping and discharging the opening spring in the driver mechanism. This actuates the interrupter to open the circuit. If the insulator above the shunt trip continues to rotate, by motor or manual actuation of the drivetrain controls, the blade opens to achieve visible isolation. The blade-hinge mechanism is actuated directly by the rotating insulator through the driver mechanism. Continued rotation of the insulator after the blade is open will "toggle" the drivetrain controls to lock the blade in its open position.

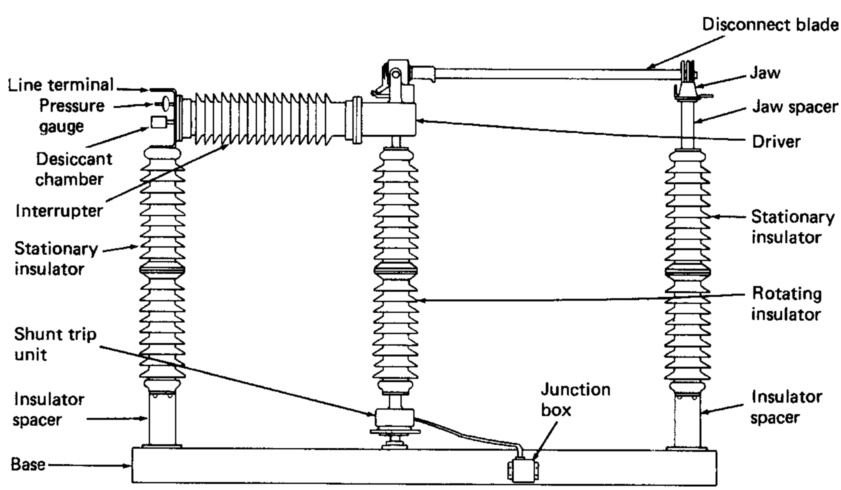


FIGURE 10-131 Single pole of blade-type circuit switcher.

For closing, the reverse rotation of the insulator first releases drivetrain toggle and allows the blade to begin closing. The shunt-trip units have already recharged during the opening operation. As the blade closes, the closing springs are charged in the driver. The last few degrees of closing rotation lock the blade in position and release the closing springs in the driver, thus closing the interrupter. The opening springs are charged as the closing springs discharge. If the unit has closed into a circuit condition that provides a trip signal to the shunt trip units, the opening process may immediately proceed since all springs are charged and all controls are ready.

The closing operation may be achieved in other designs by closing the interrupter during the opening stroke of the blade. When a close operation is called for, all that is necessary is to close the blade, because the interrupter is already closed. Because of the arc established in air for this type of closing, high-speed operation of the blade is necessary to minimize damage to contacts and prevent flashovers. Both methods of closing are proven over many years of field use.

Bladeless circuit switchers operate exactly the same as blade models, except that on opening, the insulator rotation is used only for driver and interrupter actuation. Models that depend on high-speed blade operation for closing are available in bladeless nondisconnect configuration, but circuit closing must be accomplished by other means.

For models without shunt trip, opening is accomplished by rotating the insulator to the point where the driver opening spring would normally be tripped by the shunt trip's rotation. This configuration is used where protection duty is not a function of the circuit switcher.

Transformer Protective Devices. Recently, a new transformer protective device has entered the market specifically for application on the high-side of substation transformers from 69 to 138 kV. These devices offer a 3-cycle 31,500 A interrupting rating with electronically linked pole units. They have been tested for interrupting the high TRVs from secondary-side faults and provide an economical alternative to other protective devices. They are supplied in conjunction with a separate or integrated disconnecting device on the high-voltage side of the protective device. Figure 10-132 shows a typical substation transformer protective device.

FIGURE 10-132 Substation transformer protective device. *(S&C Electric Company.)*

10.7.3 Ratings

Short-Circuit Current. Circuit switchers are "specific-duty" or "medium-fault" interrupting devices and generally fulfill a requirement between high-power fuses and circuit breakers in a transmission or distribution system. Circuit-switcher designs to date do not encompass instantaneous reclosing capability since their most common application of transformer protection precludes this requirement.

Typical short-circuit current ratings are from 4 to 40 kA on a symmetric current basis, depending on type of interrupting duty and voltage rating. The symmetric current is the highest value of the ac component in rms amperes measured at the instant of contact part.

When interruption occurs, a TRV is imposed in microseconds across the interrupting device as a result of system adjustments to the new state before a steady-state condition is achieved known as the normal-frequency recovery voltage. TRV attempts to reestablish the arc by either thermal reignitions or dielectric breakdown of the interrupting gap. Refer to ANSI C37.04 and C37.06 for TRV definitions and ratings. The TRV seen by the highside interrupting device after clearing a fault that initiated on the secondary of the transformer is high because of the characteristics of transformer.

Continuous Current. A continuous current rating is the designated limit of current in rms amperes that can be carried continuously under usual service conditions and in an ambient temperature not in excess of 40°C without exceeding temperature limits assigned to the various materials comprising the current-carrying parts or that are in contact with these parts. For further information refer to ANSI C37.04, C37.30, and IEC 694.

Circuit-switcher continuous current ratings are 1200, 1600, and 2000 A.

Short-Time Current. This is a dual rating to verify the capability of the circuit switcher to carry abnormally high currents in the closed position for short periods of time

1. Rated momentary current is the total current in asymmetric rms amperes (ac and dc components) carried for 10 cycles on a 60-Hz basis. These ratings can approach 50 times the continuous current ratings and assure the device will withstand the high electromagnetic forces from initial transient conditions of a short circuit tending to bring about mechanical damage or contact separation.
2. Rated 3-s current is 62.5% of the rated momentary current and verifies that the current-carrying parts will withstand the heating effect without excessive annealing or contact welding. Common circuit switcher momentary ratings are 61, 70, and 80 kA.

Close-and-Latch Current. For those circuit switchers that make the circuit in the interrupting device, a close-and-latch rating is required as is a close rating for the type that closes the circuit on a disconnect switch blade. This rating verifies the capability of the contact structure and operating mechanism to withstand the forces developed by closing in on a fault. Close-and-latch ratings are 30 and 40 kA. Limited-duty close ratings for the blade type are of the same magnitudes.

Rated Voltage. Rated voltages for circuit switchers are generally stated in terms of maximum system design voltages that indicate the upper limit at which the device is designed to operate. Circuit switchers are most commonly applied at 72.5 through 242 kV but are also available for special applications at 362 kV.

In order to provide safe insulation levels to ground and across the device when in the open position, the circuit switcher must withstand certain specified magnitudes and waveshapes of test voltages without flashover or puncture of any of its insulation systems. This is the rated dielectric strength and consists of

1. One-minute dry and 10-s wet low-frequency (60-Hz) withstand voltage.
2. Dry lightning-impulse (1.2×50 μs waveshape) withstand voltage. It is the positive-polarity withstand level that is the basic-insulation-level rating (BIL in kilovolts).
3. Dry and wet switching-impulse (250×2500 μs) withstand voltage. This rating applies only at 362 kV and above.
4. Dry chopped-wave impulse withstand voltage. The requirement for this rating has yet to be determined for a circuit switcher but is a standard rating for a circuit breaker. If applicable, this rating would apply to a bladeless model circuit switcher having graded gaps in series per pole.

Interrupting Time. The rated interrupting time is the maximum permissible interval between energizing the shunt trip coils at rated control voltage and the interruption of the main circuit in all poles when interrupting a current within the required interrupting capabilities. This interrupting time may be modified under certain asymmetrical current conditions as detailed in ANSI C37.04. Circuit switchers are available from 3- to 8-cycle interrupting times.

Duty Cycle. Operating duty is the short-circuit current to be interrupted, closed upon, etc. The duty cycle is a stated sequence of closing and opening operations. Various types of circuit switchers will have different duty cycles. An example is O-17s-CO where the first operation is an open followed by 17 s to permit the operating mechanism and blade to recycle in order to perform the close-open cycle.

Corona-RIV. A corona-free rating is commonly established to prevent radio and television interference. This is referred to as the radio influence voltage (RIV) and should be less than 500 μV as determined by a test circuit per NEMA Publication No. 107. As a general rule, a circuit switcher that produces no visible corona under dark conditions at 105% of rated voltage will have a negligible RIV level.

10.7.4 Selection and Application

The versatility of circuit switchers and related circuit-interrupting devices requires careful selection of the ratings, components, and accessories to be specified for a given protection and/or switching duty. The following criteria must be considered:

1. System data
   - ***a.*** Nominal service voltage
   - ***b.*** Maximum continuous current
   - ***c.*** Through-fault current withstand
   - ***d.*** TBasic impulse (insulation) level
2. Circuit-protection duty
   - ***a.*** Present and future available fault currents
   - ***b.*** Length of overhead lines or underground cables
   - ***c.*** Transformer ratings, impedance, and connections
   - ***d.*** Transient recovery voltage
   - ***e.*** Interrupting time, maximum
3. Switching duty
   - ***a.*** Inductive currents (unloaded transformers)
   - ***b.*** Small capacitive currents (unloaded lines and cables)
   - ***c.*** Inductive/capacitive currents (choke coils, capacitors, grounded or ungrounded)
   - ***d.*** Closing on fault
   - ***e.*** Closing of long lines

Approximately 80% of all circuit switchers are applied on the primary or high-voltage side of a transformer for switching and fault-protection duty. By means of protective relaying, faults which occur within the protection zone of the transformer can be detected to bring about tripping of the primary-side circuit switcher. Should the fault occur on the secondary side, which is most frequently the case, the magnitude of fault current at the circuit-switcher location is limited by the transformer impedance and is less than the available primary-fault current. Hence, the use of the term "inherent secondary-fault current," and for some circuit switchers, a dual short-circuit current rating for primary or secondary faults.

Associated with these reflected secondary faults are the inherent capacitance and inductance of the transformer windings, which generally produce higher TRVs than is the case with a primary fault. These higher TRVs can increase the difficulty of the interruption process, while conversely, the impedance-limited fault current is easier to interrupt. Circuit switchers are tested for this condition as defined in the nearly completed ANSI C37.016 which is the new standard for circuit switchers. This test standard defines secondary fault test with a TRV that is more severe than the circuit breaker test standard. Therefore, the circuit switchers demonstrate interrupting capability for this specific duty.

Primary-fault currents in excess of the switcher rating can often be cleared by source-side circuit breakers having interrupting times shorter than those of a circuit switcher. It is also common to block tripping of the circuit switcher when the fault current is in excess of its rating and pass this duty on to the source-side breaker.

Circuit switchers and interrupters are generally available in mounting arrangements similar to isolating air-disconnect switches. Since they are not freestanding, their application and that of the supporting structure design and/or optional freestanding pedestal is influenced by phase spacings desired, terminal height, altitude, atmospheric elements, seismic conditions, and wind, ice, and terminal pad loadings.

The phase spacing recommended in the ANSI C37.32 differentiates between the minimums for vertical-break disconnect switches (also applying to bus supports) and the increased spacings required for horn gap switches (switching devices utilizing arcing in air as an operating mode). From a safety viewpoint, it is important to consider the method of operation in addition to the duty applications of the particular device as a determination in selecting the phase spacing.

Elevation of the installation site, in addition to the temperature, humidity, and air-contamination characteristics, significantly affects the dielectric strength and coordination of these circuit-interrupting devices. The guidelines established in ANSI C37.30 for site elevations coupled with a further analysis of these other factors may dictate the use of higher-voltage-rated units or extra-leakage-distance porcelain standoff insulators.

Seismic, wind, ice, and terminal pad loadings must also be considered in the design of the supporting structure.

Testing. Industry standards specifically for circuit switchers have yet to be fully implemented. There is a draft standard, ANSI C37.016, which is close to being adopted. Accordingly, up to this point applicable sections of existing standards for circuit breakers and disconnect switches have been used as a guide in establishing definitions, ratings, capabilities, and test procedures for circuit switchers (see ANSI C37.09 and C37.34). In addition, specific testing for secondary fault currents and the associated higher TRVs has been completed by some circuit switcher manufacturers.

Tests can be considered to consist of four different groupings:

1. Type or design tests are for the purpose of proving the capability of the switcher to meet the specified ratings.
2. Reliability tests are part of a quality assurance program to demonstrate that the circuit-switcher design has achieved a specified mechanical reliability. They consist of growth or prototype tests to improve on the established design reliability, demonstration, or pilot run tests to assure the reliability requirements have been met, and acceptance or production tests to verify that production units comply with the demonstrated design reliability.
3. Routine or production tests are done to ensure that the production is in accordance with established procedures. This includes testing of individual components and subassemblies.
4. Installation testing to assure that the circuit switcher will perform its intended function in the system.

Accessories and Maintenance. As discussed in a preceding section, circuit-switcher manufacturers offer optional freestanding support pedestals in addition to many other accessories, some of which

10.8 AUTOMATED FEEDER SWITCHING SYSTEMS

The method known as SCADA has long been used to control transmission systems, providing the operational flexibility and speed required for efficient and reliable performance. The use of SCADA in the distribution system is becoming increasingly important as utilities move into a deregulated, competitive environment. The acronym *SCADA* is being replaced by the generic term *distribution automation* (DA), which incorporates the principle of computers operating switching and other control devices automatically in response to preprogrammed events in the system. Automated switching of distribution feeder circuits provides significant improvements in reliability, enhances operational flexibility, and increases productivity of both utility personnel and distribution lines.

Automated feeder switching systems utilize controls, sophisticated algorithms, and data communications to go beyond the reliability benefits of more traditional reclosers and automatic sectionalizers. An automated switching system can be programmed to not only reduce outage times, but do so without subjecting unfaulted circuits to "bumps" caused by intentional closing into faults required by fault-detection or troubleshooting schemes. Further, an integrated fault-isolation and restoration system can incorporate real-time evaluation of prefault loads into the algorithm to prevent overloading of a backup circuit when transferring load from a faulted circuit. This allows the utility planner more flexibility in circuit design by increasing the amount of load that can normally be carried by a given circuit.

Such systems, with current and voltage monitoring, also provide the utility with a convenient way to monitor distribution feeder circuits, thus enabling the utility to take immediate action in the event of current or voltage excursions exceeding normal operating limits. This, in turn, results in improved power quality and service for the customer.

Automated feeder switching systems can also provide the means to optimize feeder and substation loading by enabling the shifting of load from one feeder to another in a very short time frame. This same capability can yield hard-dollar cost savings associated with deferment of capital projects when coupled with planning practices that take advantage of the new technologies.

With the increasing power of computers and improving communications, it is certain that automated feeder switching systems will provide benefits that have yet to be considered in the electric utility industry.

Examples of Automated Feeder Switching Systems. A large investor-owned electric utility in New York needed to dramatically improve reliability in response to a challenge from the New York public utility commission. After analyzing its system as well as the nature and locations of faults, it was decided that an automated feeder switching system would provide the quickest route to needed improvements in electric service reliability.

A typical MES feeder "loop" is shown in Fig. 10-133. It consists of two circuits joined by a normally open tie switch. In turn, each feeder circuit will have a normally closed automated switch installed at the electrical half-way point on the circuit. All switches are equipped with voltage and current monitoring as well as communication and control units (CCUs) housing the switch controls,

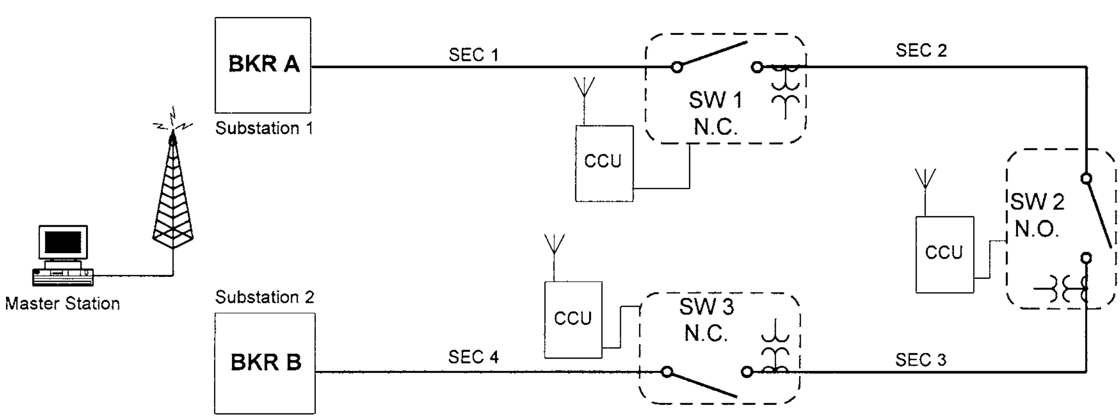


FIGURE 10-133 Typical automated feeder circuit loop.

power supplies, remote-terminal units (RTUs) and radios for establishing communication with the master station.

Faults occurring in Sec. 2 will result in a trip of substation breaker BKR A. The control at switch SW1 will sense the passage of fault current and the subsequent loss of voltage precipitated by the opening of BKR A. A sectionalizing algorithm will initiate and open the switch after a predetermined time delay set to coordinate with the recloser interval of BKR A. The station breaker is reclosed to determine if the fault condition is temporary or permanent. A typical reclosing breaker may reclose up to 3 times on some systems. Once SW1 is open and the fault is isolated, BKR A will reclose, restoring service to the customers in Sec. 1. At this point, the crew can proceed directly to SW1 and commence the search for the fault from there, further reducing the amount of time for fault location.

Faults occurring in Section 1 will result in a lockout of substation BKR A. At this point, SW1 will see a loss of voltage and transmit an alarm to the master station. The master will then interrogate SW1 and find that it saw loss of voltage, but no overcurrent. This condition indicates a section 1 fault. The master station will then initiate the section 1 fault algorithm and interrogate SW1 to determine the load in section 2 just prior to the fault. The master then interrogates SW3 for load information. If the addition of section 2 load to the current load on section 3 will not result in an overload condition, then the master station will instruct SW2 to close, restoring service to section 2.

The MES system can be operated in three modes. The first is traditional SCADA, where the dispatcher must identify and analyze circuit conditions before deciding on restoration procedures, which must then be sent to the field devices individually. In the second mode, the master station computer will identify and analyze the circuit condition and make an operational recommendation to the dispatcher, who can then choose to accept the recommendation and execute a simple acknowledge command which will automatically issue appropriate commands to field devices. The third mode identifies and analyzes circuit conditions and automatically issues commands to field devices to isolate faulted sections and restore service to unfaulted sections. It should be noted that the system can utilize up to seven switching devices on a given circuit loop similar to that shown in Fig. 10-133. The figure was simplified for clarity in this example.

A large metropolitan utility serving customers in and around Chicago wanted to improve electric service. However, they did not have a SCADA master station at the beginning of the project and needed a way to improve service without installing one. In this case, the utility installed a system which utilizes distributed intelligence in the switch controllers to do all analysis as well as fault isolation and circuit restoration. Such a system operates in much the same way as described in the example above. The key difference is that this system does not utilize the first two modes of operation: traditional SCADA and suggesting operational recommendations. The advantage is the ability to install the system without the need for a communications network or SCADA master station. This system does its work automatically without human intervention—or the need for a communication system. It should be noted that the system can be easily disabled in the field for line work if desired and can be upgraded to work with a master control station.

10.8.1 Automated Switches

A key part of an automated feeder switching system is the automated switch. The term “automated” in this context means the switch is designed for use on an automated or SCADA system. In order to be automated, existing switches may be retrofitted with motor operators, current and voltage sensors, RTUs and communication devices to allow the remote operation necessary to realize the benefits available with automated feeder switching systems. However, switches designed for occasional, manual operation may not be entirely suitable for operation on an automated distribution circuit feeder. Manual switches are typically not designed to be operated the hundreds of times required by a fully automated system over the life of a typical switch. Nor are they ordinarily designed for duty-cycle fault-closing to allow the system operator to inadvertently close into a fault from the SCADA master station—and still leave the switch in an operable condition.

More recently, switches designed specifically for automation have appeared in the market (Fig. 10-134). Such switches incorporate design features that make them particularly applicable for use in an automated feeder switching system:

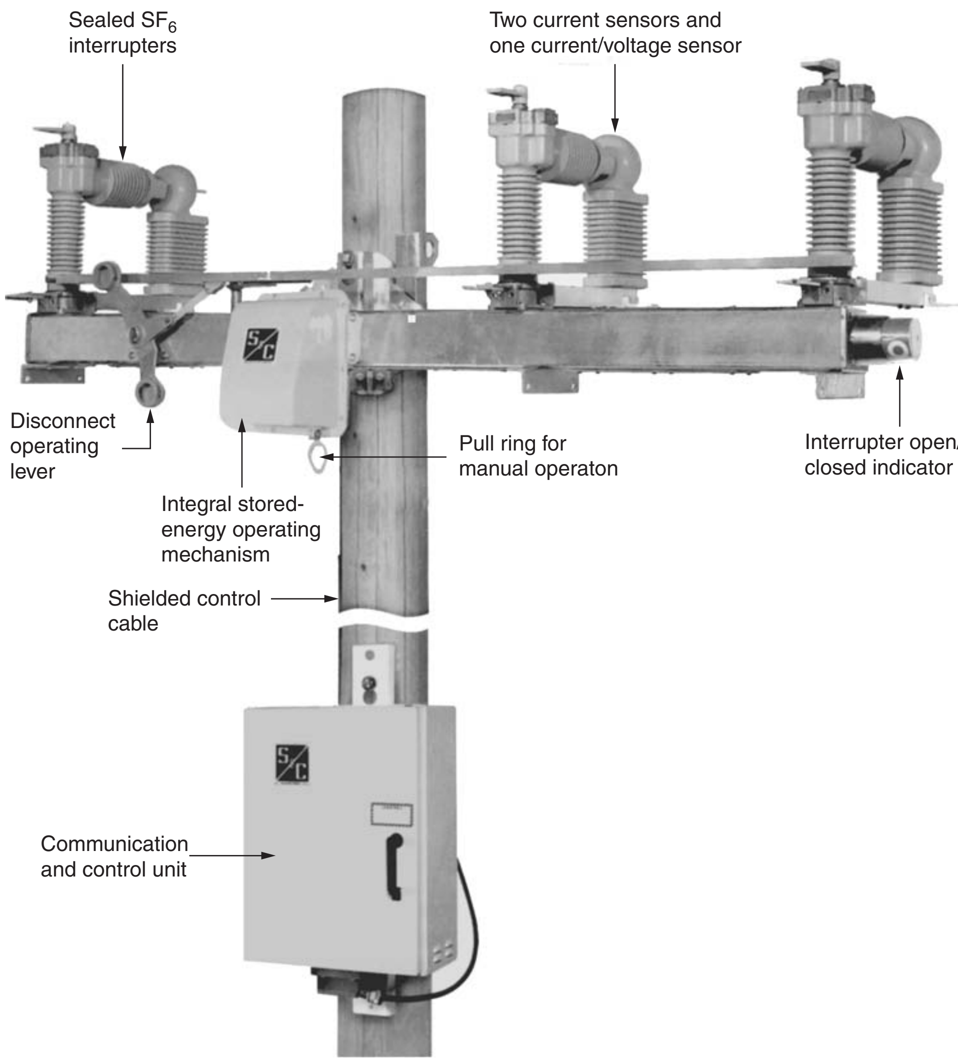


FIGURE 10-134 Scada-Mate switching system. *(S&C Electric Company.)*

1. *Duty-cycle fault-closing* allows the switch to be closed into a typical fault several times before experiencing damage severe enough to render the switch inoperable.
2. *Integrated voltage and current sensors* provide the ability to monitor voltages, currents, and loads that are in turn used as inputs to algorithms to effect automated switching for fault isolation and restoration and for shifting loads for circuit optimization.
3. *Integrated operating mechanisms* enable the switches to be operated remotely via computer commands. Integration with the switch ensures optimum operation without the need for cumbersome ground-to-switch linkages.
4. *Integrated load interrupters* should be designed to allow operation under any weather conditions since it will not be possible to visibly inspect the switch for ice or other problems prior to operation.

5. *Integrated control power sources* eliminate the need to rely on locally available control power sources—or to install such power sources.
6. *Integrated visible air-gap isolation* provides the visible air gap when needed for certain types of line work.

In addition, an associated control package should include switch-operating controls, a local/remote switch, backup power for dead-line SCADA operation, a remote-terminal unit, and data-communication devices. The entire package should be assembled and tested for proper operation by a single supplier to eliminate the need for the utility to perform the integration. The control box should be separately located from the switch to allow access by technicians who are not qualified in high-voltage operations. In underground switchgear applications, the control should be isolated from the high-voltage compartments of the switchgear.

SECTION 11
ALTERNATE SOURCES OF POWER

Mahan V. Aware
Electrical Engineering Department, Visvesvaraya National Institute Technology, Nagpur, India

Ramesh C. Bansal
Electrical & Electronics Engineering Department, Birla Institute of Technology & Science, Pilani (Rajasthan), India

Paul Butler
Sandia National Laboratory, Albuquerque, NM

Palmer W. Carlin
National Renewable Energy Laboratory, Golden, CO

Raymond Fortuna
U.S. Department of Energy, Washington, DC

Neil B. Morley
Mechanical and Aerospace Engineering Department, University of California, Los Angeles, CA

Mark. S. Tillack
Mechanical and Aerospace Engineering Departyment

Robert D. Weaver
Consultant, Auburn, CA

Ahmed F. Zobaa
Electrical Power & Machines Department, Cairo University, Cairo, Egypt

Charles P. (Sandy) Butterfield
National Renewable Energy Laboratory, Golden, CO

Michael Milligan
National Renewable Energy Laboratory, Golden, CO

Eduard Muljadi
National Renewable Energy Laboratory, Golden, CO

Karin Sinclair
National Renewable Energy Laboratory, Golden, CO

Yih-Huei Wan
National Renewable Energy Laboratory, Golden, CO

CONTENTS

11.1 TRADITIONAL ENERGY SOURCES

BY RAMESH BANSAL

Electricity has been generated for the purpose of powering human needs for more than 120 years from various energy sources. The importance of dependable electricity generation was revealed when it became apparent that electricity was useful for providing heat, light, and power for human activities. Today, traditional (conventional) sources of power generation are fossil fuels (coal, petroleum, natural gas), hydro and nuclear power systems.

11.1.1 Coal

The primary sources for coal are China, the United States, and the former USSR. These countries have 75% of the coal reserves. Coal was the first fossil fuel used for producing electricity. Electricity is produced at a coal-fired fossil plant by the process of heating water in a boiler to about 540°C to produce steam. The steam, under tremendous pressure of about 130 Kg/cm^2, flows into a turbine, which spins a generator to produce electricity. Many environmental problems are associated with the use of coal including nitrous oxides, sulfur oxides, CO_2, as well as particulate matter is released when coal is burned. Nitrous oxides and sulfur oxides cause acid rain and CO_2 is responsible for global warming.

11.1.2 Crude Oil (Petroleum)

The use of oil started after the 1860s. Oil is liquid hydrocarbon. There are three main products of oil: kerosene, fuel oil, and gasoline. About 75% of the oil production is controlled by OPEC (Organization of Petroleum Exporting Countries), which was formed in 1973. OPEC can negotiate as a block to sell their oil, and thus are able to set higher prices. The advantages of oil are easy to handle, store, and transport. Also, it is used for things other than energy: productions of plastic, lubrication, etc. While oil burns cleaner than coal, the same pollutants are produced and the same environmental problems are associated with oil production, transport, and combustion.

11.1.3 Natural Gas

Natural gas is available in three types: methane, propane, and butane. Natural gas is often found in association with crude oil. The largest reserves are found in the Persian Gulf countries. Advantages of natural gas are that it burns cleaner than the other fossil fuels, less nitrous oxides and sulfur oxides are produced, and less particulate matter, but CO_2 is still produced. Combustion turbines can run on natural gas or low-sulfur fuel oil and are designed to start quickly to meet the demand for electricity during peak operating periods. Air enters at the front of the unit and is compressed, mixed with natural gas or oil, and ignited. The hot gas then expands through turbine blades to turn the generator and produce electricity.

11.1.4 Hydro

Hydroelectric power is a form of power that utilizes the energy released by water falling on the turbine blades that rotates the generator to produce electricity. Hydroelectricity is a renewable energy source, since the water that flows in rivers has come from precipitation such as rain or snow. The energy that may be extracted from water depends not only on the volume but also on head (the difference in height between the water crest [or source] and the water outflow).

TABLE 11-1 Installed Hydroelectric Capacity

Country	Hydroelectric installed capacity (MW)
United States	79,511
Canada	66,954
China	65,000
Brazil	57,517
Russia	44,000
Norway	27,528

Hydroelectric power supplies about 20% of the world's electricity. Norway produces virtually all of its electricity from hydro, while Iceland and Austria produce 83% and 67%, respectively, of their electricity from hydro. Countries with their hydroelectric installed capacities are shown in Table 11-1.

The world's largest hydroelectric power stations in operation are at Itaipu, Brazil with an installed capacity of 12,600 MW and Three Gorges Dam, China with installed capacity of 18,200 MW, which is scheduled to be completed by 2009.

The chief advantage of hydro systems is elimination of the cost of fuel. Hydroelectric plants are immune to price increases for fossil fuels and do not require imported fuel. Hydroelectric plants tend to have longer lives than fossil-fuel-fired generation, with life spans of 50 to 100 years. Operation and maintenance costs also tend to be low since plants are generally heavily automated and have fewer personnel on-site during normal operation. Hydroelectric plants are pollution free. Since the generating units can be started and stopped quickly, they can follow system loads efficiently, and may be able to reshape water flows to more closely match daily and seasonal system energy demands. Concerns have been raised by environmentalists that large hydroelectric projects might be disruptive to surrounding aquatic ecosystems. Another disadvantage of hydroelectric dams is the need to relocate the people living where the reservoirs are planned.

11.1.5 Nuclear Fuel

The first commercial nuclear power stations started operation in the 1950s. There are now some 440 commercial nuclear power reactors operating in 31 countries, with over 364,000 MW of total capacity. Main countries producing electricity from nuclear power are the United States, France, and

Japan, having installed capacities of 97,585, 63,474, and 46,343 MW, respectively, as of May 2005. Nuclear fission, or the splitting of atoms of ^{235}U, releases energy, which is used to heat water and produce steam, which drives turbines to produce electricity.

Although uranium is mined, very little of it is needed compared to coal or oil, so the mining itself is not as big an environmental concern as it is for the fossil fuels. Advantages of nuclear power include the lack of pollution-causing emissions. There are great dangers associated with nuclear power production, however, since a by-product of the process, plutonium, can be used to make nuclear weapons. In addition, there are problems associated with how to dispose off nuclear waste and how to deal with decommissioning old power plants that are no longer productive.

The two most common types of nuclear power plants are boiling water reactor (BWR) and pressurized water reactor (PWR), both using water as coolant. In BWR plants, the water is allowed to boil in the reactor core, the steam is then passed through the turbine, which runs the generator to produce electricity. In PWR plants, there are two more cycles linked by the heat exchanger. In PWR, fuel rods are placed in the reactor vessel to make up the core—the part of the plant that produces heat. When a uranium atom splits in the process called nuclear fission, it gives off energy in the form of heat. To regulate the heat-producing process, control rods and borated water are used. The borated water speeds up or slows down the fission process, and the control rods shut down the reaction when they are inserted between the fuel rods. A nuclear plant works in much the same way that a dam or fossil fuel plant does, in which large turbine blades are used to run a generator to produce electricity. At a hydroelectric dam, the force of the falling water spins the turbine blades, while at a coal-fired or nuclear plant, the force of steam spins the blades. A nuclear plant, however, uses uranium instead of coal as a fuel to make steam.

11.2 RENEWABLE ENERGY TECHNOLOGIES

BY RAMESH BANSAL

The energy crisis, which began in 1973, caused petroleum supplies to decrease and prices to rise exorbitantly. This crisis forced developing countries to reduce or postpone important development programs, so they could purchase petroleum to keep their economies operating. It created the urgent necessity to find and develop alternative energy sources, as other fossil fuels (coal, oil, and natural gas), nuclear energy, and renewable energy resources.

There are concerns about nuclear energy because of the associated accident risks; waste disposal difficulties, nuclear terrorism, and nuclear weapon proliferation are dangerous in themselves. Acquiring nuclear energy from the industrialized world could, moreover, result in greater technological and economic dependence on developed countries. World's proved fossil fuel resources might be exhausted in about 100 years, thus making situation alarming. A more feasible alternative to petroleum, coal, and nuclear reactors in developing countries is the *direct* and *indirect* use of solar energy, which is renewable, abundant, decentralized, and nonpolluting.

Each day, the sun sends to earth many thousands of times more energy than we attain from other sources (the equivalent of 200 times the energy consumed by the United States in 1 year). This energy can be captured *directly* as radiation or—even more significantly—*indirectly* in waterfalls, wind, and green plants. Taking into account that the technology needed for exploiting renewable energy resources is simple and relatively economical, it is important from a strategic point of view that energy planning in Third World countries, particularly in the humid tropics, be oriented to developing the solar alternative. It offers them one of the few opportunities to develop independently of the industrialized countries. This section briefly describes various renewable energy sources, that is, solar, wind, small hydro, biomass, geothermal, tidal, magnetohydrodynamic (MHD), and ocean thermal energy conversion (OTEC).

11.2.1 Solar Energy

Solar power describes a number of methods of harnessing energy from the light of the sun. It is already in widespread use where other supplies of power are absent such as in remote locations and in space. As the earth orbits the sun, it receives approximately 1,020 W/m^2 at sea level.

Solar power may be classified as direct and indirect. Direct solar power involves only one transformation into a usable form, for example, sunlight hits a photovoltaic cell to create electricity and warms the surface or heats the water when the light is converted to heat by interacting with matter. Indirect solar power involves more than one transformation to reach a usable form. Many other types of power generation are indirectly solar-powered, for example, (i) vegetation use photosynthesis to convert solar energy to chemical energy, which can later be burned as fuel to generate electricity; (ii) energy obtained from oil, coal, and peat originated as solar energy captured by vegetation in the remote geological past and fossilised; (iii) hydroelectric dams and wind turbines are indirectly powered by solar energy through its interaction with the earth's atmosphere and the resulting weather phenomena; (iv) energy obtained from methane (natural gas) may be derived from solar energy either as a biofuel or fossil fuel; (v) ocean thermal energy production uses the thermal and gradients that are present across ocean depths to generate power.

Solar power can also be classified as passive or active. *Passive solar systems* are systems that do not involve the input of any other forms of energy apart from the incoming sunlight. *Active solar systems* are those that use additional mechanisms such as circulation pumps, air blowers, or automatic systems that aim collectors at the sun. Effective use of solar radiation often requires the radiation (light) to be focused to give a higher intensity beam, that is, parabolic dish, parabolic trough, etc., are used to concentrate light at a point or a line. At the focus, high-concentration photovoltaic cells (solar cells) or a thermal energy "receiver" may be placed. Most of the solar energy used today is harnessed as heat or electricity. Solar design aims the use of architectural features to replace the use of grid electricity and fossil fuels with the use of solar energy and decrease the energy needed in a home or building with insulation and efficient lighting and appliances. Following are the main applications of solar energy:

Photovoltaic systems: Solar cells, also known as photovoltaic cells, use the photovoltaic effect of semiconductors to generate electricity directly from the sunlight. Because of high manufacturing costs, their use has been in limited until recently. One cost-effective use has been in very low-power devices such as calculators with LCDs. Another use has been in remote applications such as roadside emergency telephones, remote sensing, cathodic protection of pipelines, and limited to isolated home power applications. A third use has been to power orbiting satellites and other spacecraft. However, the continual decline of manufacturing costs (dropping at 3% to 5% a year in recent years) is expanding the range of cost-effective uses.

Solar heating: Solar hot water systems are quite common in some countries where a small flat panel collector is mounted on the roof and is able to meet most of a household's hot water needs. Cheaper flat panel collectors are also often used to heat swimming pools, thereby extending the swimming season. There are some new applications of thermal hot water, like air cooling, currently under development.

Solar cooker: Taps the sun's power in an insulated box, which has been successfully used for cooking. Solar cooking is helping many developing countries both by reducing the demands for local firewood and maintaining a cleaner environment for the cooks.

11.2.2 Wind Energy

Among the renewable sources of energy available today for generation of electrical power, wind energy stands foremost because of the no pollution, relatively low capital cost involved and the short gestation period required. Wind-powered systems have been widely used since the tenth century for water pumping, grinding grain, and other low-power applications. There were several early attempts to build large-scale wind-powered systems to generate electricity. Recently, wind turbine of 4.5 MW and rotor diameter of more than 112.8 m has been in operation.

Today, wind energy is the fastest growing energy source. Presently, wind power meets the electricity needs of more than 35 million people. Globally, the wind power industry employs around 70,000 people and is worth more than $5 billion. According to Global Wind Energy Council (GWEC), global wind power capacity has increased from 7,600 MW at the end of 1997 to 47,337 MW by February 2005. Main countries producing electricity from wind are Germany, Spain, the United States, Denmark, and India, having their installed capacities 16,629, 8,263, 6,470, 3,117, and 3,000 MW, respectively, by February 2005. Today, wind power accounts for about 0.4% of the world's electricity demand. An analysis by the European Wind Energy Association (EWEA) shows that there are no technical, economic, or resource limitations that prevent wind power from developing to nearly 12% of the world's electricity supply by 2020, but with strong political commitment worldwide, the wind energy industry could install an estimated 1200,000 MW by 2020.

The total wind power P_w that is available to a wind turbine is given by

$$P_w = (\rho A V^3)/2 \tag{11-1}$$

where ρ is the density of the air in kg/m^3, A is the exposed area in m^2, and V is the velocity in m/s. The maximum power that can be realized from a wind system is 59.3% of the total wind power. The power in the wind is converted to mechanical power with an efficiency (coefficient of performance) c_p, which is transmitted to the generator through a mechanical transmission with efficiency n_m, and is converted to electricity with an efficiency n_g. The electrical power output is then

$$P_e = c_p\, n_m\, n_g\, P_w \tag{11-2}$$

For a given system, P_w and P_e will vary with wind speed. As the wind increases from a low value, the turbine is able to overcome all mechanical and electrical losses and start delivering electrical power to the load at cut-in speed V_C. The rated power output of the generator is reached at rated wind speed V_R. Above V_R, some wind power is spilled to maintain constant power output. At the furling speed V_F, the machine is shut down to protect it from high winds. Seasonal and diurnal variation has significant effect on wind. Other factor, which affects power from wind, is height of wind turbine. Wind speed increases with the height because of friction at the earth's surface.

There are a number of ways of classifying wind systems, for example, according to size of power output, rotational speed of wind turbines, orientation of wind turbines, etc. According to the size of power output, wind systems may be classified, for example, as small, medium, and large.

According to the rotational speed, wind turbines are classified as fixed speed and adjustable speed generators. In fixed speed generators, the rotor is held constant by continuously adjusting the blade pitch and/or generator characteristics. For synchronous generators, the requirement of constant speed is very rigid and only minor fluctuations of about 1% for short durations could be allowed. As the wind fluctuates, a control mechanism becomes necessary to vary the pitch of the rotor so that the power derived from the wind system is held fairly constant. Induction generators with small negative slip can also be considered as constant speed. Induction generators are simpler than synchronous generators. They are easier to operate, control and maintain, have no synchronization problem, and are economical. Modern high-power wind turbines are capable of adjustable speed operation. Key advantages of adjustable speed generators compared to fixed speed generators are that they are cost effective, provide simple pitch control, and yield higher output for both low and high wind speeds.

According to the orientation of turbines, wind turbines are classified as horizontal and vertical axis machines. In horizontal axis wind turbines (HAWT), the axis of rotation is parallel to the direction of the wind. Depending upon the number of blades these may be classified as single-bladed, double-bladed, three-bladed, multibladed, and bicycle-bladed. In vertical axis wind turbines (VAWT), the axis of rotation is perpendicular to the direction of wind. These machines are also called *crosswind axis machines*. Main designs of vertical axis machines are Savonious and Darrieus rotors. The principal advantages of VAWT over conventional HAWT are that VAWT are omnidirectional, that is, they accept the wind from any direction. The vertical axis rotation also permits mounting the generator and gear at the ground level. On the negative side VAWT requires guy wires attached to the top for the support, which may limit its application particularly for the offshore sites.

Wind speeds over the open ocean average 30% to 50% higher than on land, while turbulence is reduced because of the absence of surface obstructions such as hills, trees, and buildings. Since power increases as the cube of the wind speed, the substantial increases in output that can be achieved in offshore wind power systems can more than offset the increased cost of sitting wind turbine in water. Offshore designs are therefore made to meet different requirements. Wind machines must have marine grade components, seals, and coatings to protect them from corrosion. In Europe, offshore projects are now springing up off the coasts of Denmark, Sweden, the United Kingdom, France, Germany, Belgium, Irelands, the Netherlands, and Scotland.

11.2.3 Small Hydropower

Although this technology is not new, its wide application to small waterfalls and other potential sites is new. It is best suited to high falls with low volume, such as occur in high valleys in the mountains. It is the application of hydroelectric power on a commercial scale serving a small community. These plants are classified by power and size of waterfall. A generating capacity of up to 10 MW is becoming generally accepted as the upper limit of small hydro, although this may be stretched up to 30 MW in some countries. Small hydro can be further subdivided into mini-hydro, usually defined as less than 1,000 kW, and micro-hydro which is less than 100 kW.

Hydroelectric power is the technology of generating electric power from the movement of water through rivers, streams, and tides. Water is fed via a channel to a turbine where it strikes the turbine blades and causes the shaft to rotate. To generate electricity, the rotating shaft is connected to a generator which converts the motion of the shaft into electrical energy. Small hydro is often developed using existing dams or through development of new dams whose primary purpose is river and lake water-level control, or irrigation. A small-scale hydroelectric facility requires a sizeable flow of water and a reasonable height of fall of water, called the *head*. Another advantage of using water resources is that hydraulic works can be made simple, and large constructions, such as dams, are not usually required. When dams are necessary, they affect less area than in lower zones because of the steepness of the terrain. Dams, which exploit the kinetic energy of water by raising small quantities of water to heights through the use of regulated pressure valves, can provide water for domestic uses and for agriculture in areas that are moderately higher than adjacent water courses. Another interesting possibility is the utilization of induction generators for supplementing small hydroelectric plants, which require lower initial costs and have technical operation advantages over synchronous generators.

11.2.4 Biomass Energy

Biomass contributes 14% of the world's primary energy demands, and in developing countries it constitutes 35% of the primary energy supply. Biomass is an energy carrier that can be used in solid, liquid, and gaseous forms, and is a versatile source of energy that can produce electricity, heat, transport fuel, and can be stored conveniently. Energy production of biomass units ranges from small scale to multi-megawatt size. The main biomass conversion technologies are biomass gasifiers and biogas generation.

Biomass is organic nonfossil material. In other words, biomass is all plant, trees, and animal matter on the earth's surface. Humans, domestic animals, and crops comprise somewhere between 40% to 60% of the earth's biomass. In many ways biomass can be considered as a form of stored solar energy. The energy of the sun is "captured" through the process of photosynthesis in growing plants. Biomass is sometimes burned as fuel for cooking and to produce electricity and heat. Methanol and ethanol are popular sources of alternative energy produced by the fermentation of organic matter, such as manure, under anaerobic conditions. The use of biogas is encouraged because methane burns with a clean flame and produces little pollution. Digestion of manure occurs in a digester, which must be strong enough to withstand the buildup of pressure and must provide anaerobic conditions for the bacteria inside.

11.2.5 Geothermal Energy

Electricity from geothermal energy is generated by utilizing naturally occurring geological heat sources. Geothermal-generated electricity was first produced at Larderello, Italy, in 1904. Since then, the use of geothermal energy for electricity has grown worldwide to about 8,000 MW of which the United States produces 2,700 MW. The largest dry steam field in the world is The Geysers, about 90 miles north of San Francisco, began in 1960, which produces 2,000 MW. Geothermal power is generated in over 20 countries around the world including Iceland (producing 17% of its electricity from geothermal sources), the United States, Italy, France, New Zealand, Mexico, Philippines, Indonesia, and Japan.

Large scale electrical generation is possible in areas near geysers or hot springs by utilizing naturally occurring steam, superheated ground water, or using geothermal heat to heat a heat-transfer fluid. Experiments are in progress to make deep wells into hot dry rocks (HDR), which can be economically used to heat water pumped down from the surface. Geothermal areas without steam are called HDR. HDR programs are currently being developed in Australia, France, Switzerland, and Germany. Magma (molten rock) resources offer extremely high-temperature geothermal opportunities, but existing technology does not allow recovery of heat from these resources.

Although geothermal sites are capable of providing heat for many decades, eventually they are depleted as the ground cools. It can be said that the geothermal resource is not strictly renewable in the same sense as the hydro resource. Currently, there are few geothermal resource areas capable of generating electricity at a cost competitive with other energy sources, such as natural gas and coal. Some do not have a high enough temperature to produce steam and others don't have the water to produce steam, which is necessary for current plant designs. Also, instead of producing electricity, lower temperature areas can provide space and process heating.

11.2.6 Tidal Energy

Tidal power is a means of electricity generation achieved by capturing the energy contained in moving water mass due to tides. Two types of tidal energy can be extracted: kinetic energy of currents due to the tides and potential energy from the difference in height (or *head*) between high and low tides. The extraction of potential energy involves building a barrage. The barrage traps a water level inside a basin. Head is created when the water level outside of the basin changes relative to the water level inside. The head is used to drive turbines. In any design this leads to a decrease of tidal range inside the basin, implying a reduced transfer of water between the basin and the sea. This reduced transfer of water accounts for the energy produced by the scheme.

Tidal power is classified as a renewable energy source, because tides are caused by the orbital mechanics of the solar system and are considered inexhaustible within a human time frame. The root source of the energy comes from the slow deceleration of the earth's rotation. The moon gains energy from this interaction and is slowly receding from the earth. Tidal power has great potential for future power and electricity generation because of the total amount of energy contained in this rotation. The efficiency of tidal power generation largely depends on the amplitude of the tidal swell, which can be up to 10 m where the periodic tidal waves funnel into rivers and fjords. Selection of location is critical for a tidal power generator. The potential energy contained in a volume of water is

$$E = xMg \tag{11-3}$$

where x is the height of the tide, M is the mass of water, and g is the acceleration due to gravity. Therefore, a tidal energy generator must be placed in a location with very high-amplitude tides. Suitable locations have been found in the former USSR, the United States, Canada, Australia, Korea, the United Kingdom and in many other countries.

11.2.7 Magnetohydrodynamic Generation

MHD power generation is a method of direct conversion of heat into electrical energy. Kinetic energy of the fluid is converted into electrical power by the interaction of the electrical conducting fluid under the influence of magnetic field. In thermal generation of electric energy, the heat released by the fuel is converted to rotational mechanical energy by means of a thermocycle. The mechanical energy is then used to rotate the electric generator. Thus, two stages of energy conversion are involved in which the heat to mechanical energy conversion has inherently low efficiency. Also, the rotating machine has its associated losses and maintenance problems. In MHD technology, the hot gases produced by the combustion of fuel without the need for mechanical moving parts directly generate electric energy.

The fluid conductor is typically an ionized flue gas resulting from combustion of coal or other fossil fuels. The conductive fluid flows through the magnetic field, inducing an electric field by the Faraday effect. The electric field is orthogonal to both the fluid velocity and magnetic field vectors. As a result, potential difference is developed between the two walls of the duct. The direct current generated is converted to alternating current by a solid-state inverter. A typical MHD plant requires combustion gases of about 2,650ºC and a pressure of 500 to 1,000 kPa. Commercial scale MHD plants will use superconducting magnets. To achieve superconducting properties, the magnets must be cooled to around 4K. The MHD technology is in the relatively early development stage, although test data indicate that there are no fundamental barriers for commercialization of MHD technology. Several prototype units are being tested in the United States.

11.2.8 Ocean Thermal Energy

OTEC is an energy technology that converts solar radiation to electrical power. OTEC systems use the ocean's natural thermal gradient—the fact that the ocean's layers of water have different temperatures to drive a power-producing cycle. As long as the temperature between the warm surface water and the cold deep water differs by about 20°C, an OTEC system can produce a significant amount of power. The oceans are thus a vast renewable resource, with the potential to help us produce billions of watts of electrical power. The potential is estimated to be about 10^{13} W of base load power generation. The cold, deep seawater used in the OTEC process is also rich in nutrients, and it can be used to culture both marine organisms and plant life near the shore or on land.

The main advantages of OTEC are that (i) it uses clean, renewable, and natural resources, (ii) warm surface seawater and cold water from the ocean depths replace fossil fuels to produce electricity, (iii) suitably designed OTEC plants produce negligible pollution, and (iv) it can produce freshwater as well as electricity, which is a significant advantage in island areas where freshwater is limited. The disadvantages of OTEC are (i) OTEC-produced electricity at present costs more than the electricity generated from fossil fuels at their current costs, (ii) plants must be located where a difference of about 20ºC occurs year-round, (iii) ocean depths must be available fairly close to shore-based facilities for economic operation, and (iv) no energy company may put money in this project because it only had been tested in a very small scale.

OTEC covers 71% of the earth's surface and acts as a natural collector and store of solar energy. On an average day, 60 million km^2 of tropical seas absorb an amount of solar radiation equivalent in heat content to about 245 billion bbl of oil. The main countries in which OTEC plants exist are the United States with installed capacity of 100 MW, the United Kingdom, the Netherlands, Japan, and Taiwan with capacity of about 10 MW. By 2010, about 1,000 OTEC plants are expected to be installed in the range 1 to 100 MW to generate about 50,000 MW.

BIBLIOGRAPHY

T. S. Bhatti, R. C. Bansal, and D. P. Kothari (Eds.), "*Small Hydro Power Systems*," Dhanpat Rai & Sons, Delhi, India, 2004.

R. C. Bansal, T. S. Bhatti, and D. P. Kothari, "On Some of the Design Aspects of Wind Energy Conversion Systems," *Int. J. of Energy Conversion & Manage.*, vol. 43, no. 16, pp. 2175–2187, Nov. 2002.

M. H. Dickson and M. Fanelli (Eds.), "*Geothermal Energy,*" John Wiley & Sons, Chichester, England, 1995.

T. Jiandong, Z. Naibo, W. Xianhuan, H. Jing, and D. Huishen (Eds.), "*Mini Hydropower,*" John Wiley & Sons, Chichester, England, 1995.

R. Hunter and G. Elliot, *"Wind-Diesel Systems, A Guide to the Technology and its Implementation,"* Cambridge University Press, Great Britain, 1994.

H. Nacfaire (Ed.), "*Wind-Diesel and Wind Autonomous Energy Systems,*" Elsevier Applied Science, London, 1989.

J. W. Twidel and A. D. Weir, "*Renewable Energy Sources,*" English Language Book Society (ELBS), Cambridge University Press, Great Britain, 1986.

P. Gipe, "*Wind Power*," Chelsea Green Publishing Company, Post Mills, VT, 1995.

R. W. Thresher and D. M. Dodge, "Trends in the Evolution of Wind Turbine Generator Configurations and Systems," *Int. J. Wind Energy*, vol. 1, pp. 70–85, 1998.

D. C. Quarton, "The Evolution of Wind Turbine Design Analysis—A Twenty Years Progress Review," *Int. J. Wind Energy*, vol. 1, pp. 5–24, 1998.

11.3 SOLAR ENERGY

By AHMED ZOBAA, CAIRO (Egypt) University

The sun's energy arrives on earth in the primary form of heat and light. Other aspects of solar radiation are less easily perceived and their detection often requires sophisticated equipment. All solar radiation travels through space in waves, and it is the length of these waves (the shortest is less than a millionth of an inch, the longest more than a thousand yards) by which all solar radiation is classified. The aggregate of all radiation aspects of the sun is called the *solar spectrum*.

There are two important facets about the solar spectrum:

1. While the sun emits radiation in all wavelengths, it is the short wavelength radiation that accounts for the majority of energy in the solar spectrum. For example, the portion of the spectrum perceived as the visible light is a relatively small segment compared to the variety of spectrum wavelengths, yet accounts for 46% of the energy radiating from the sun. Another 49%, which is perceived as heat, is derived from the infrared band of the spectrum.
2. The proportion of different wavelengths in the solar spectrum does not change and therefore the energy output of the sun remains constant. A measurement of this phenomenon is known as the *solar constant*, defined as the amount of heat energy delivered by solar radiation to a square foot of material set perpendicular to the sun's rays for 1 h at the outer edge of the earth's atmosphere. The solar constant measurement is about 429.2 Btu with minimal changes over the year. The energy measured as the solar constant is not a measure of the amount of solar energy that actually reaches the earth's surface, since as much as 35% of all the solar radiation intercepted by the earth and its surrounding atmosphere is reflected back into space. Additionally, water vapor and atmospheric gases absorb another 15%. As a global average only about 35% to 40% of the solar radiation entering the atmosphere actually reaches the earth's surface.

11.3.1 Solar Constant

The solar constant is the amount of energy received at the top of the earth's atmosphere on a surface oriented perpendicular to the sun's rays (at the mean distance of the earth from the sun). The generally accepted solar constant of 1,368 W/m^2 is a satellite measured yearly average.

In order to calculate the solar constant, the following equation is used:

$$S = E\,(\text{sun}) \times (R\,(\text{sun})/r)^2 \tag{11-4}$$

where S = solar constant
E = surface irradiance of the sun
R = 6.96 × 105 km = radius of the sun
r = 1.5 × 108 km = average sun-earth distance

The solar constant is not really constant; this is due to the variation of the intensity of the sun due to sunspots. *Sunspots* are convective activity in the upper layer of the sun. Their number varies in cycles of 22 years and influences solar luminosity. Another reason why S varies is because of changes in the average distance earth-sun "eccentricity." The eccentricity varies regularly with periods of about 100,000 and 400,000 years. The maximum change in S associated with variation in eccentricity is about 0.1%. The angle of tilt of the earth's axis of rotation varies between 22° and 24.5° with a periodicity of about 40,000 years.

Table 11-2 presents some values of solar constants in kWh/m^2/day for different countries in the world.

11.3.2 Radiation Received at Earth's Surface

The earth gets only 2 billionths of the sun's energy, but that is still a lot. However, life (through photosynthesis) uses only 0.023% of the energy that reaches the surface of the earth (Fig. 11-1).

TABLE 11-2 Solar Constants in kWh/m²/day for Different Countries in the World

Country	State/City	Latitude	Longitude	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year avg
DZ	Alger	36′ 50″ N	3′ E	2.22	2.94	3.87	5	5.88	6.69	7.23	6.48	5.15	3.53	2.43	2.02	4.45
EG	Cairo	29′ 35″ N	31′ 09″ E	3.39	4.17	5.24	6.49	7.11	8	7.88	7.4	6.42	5.07	3.86	3.19	5.68
GM	Gambia	13′ 28″ N	16′ 39″ W	5.01	5.93	6.67	6.92	6.72	6.06	5.48	5.09	5.3	5.67	5.36	4.95	5.76
LR	Liberia	6′ 30″ N	9′ 30″ W	5.43	5.72	5.59	5.31	5.11	4.61	4.25	4.19	4.67	5.13	5.24	5.2	5.03
KE	Nairobi	1′ 16″ S	36′ 48″ E	6.05	6.24	6.07	5.7	5.42	5.14	4.88	5.09	5.78	6.03	5.48	5.6	5.62
CA	San Francisco	38′ 31″ N	121′ 30″ W	2.35	3.33	4.42	5.95	6.84	7.39	7.55	6.51	5.75	3.92	2.65	2.06	4.89
CO	Denver	39′ 45″ N	104′ 52″ W	2.25	3.2	4.32	5.61	6.11	6.71	6.5	5.86	5.47	4.01	2.59	1.98	4.55
CT	Hartford	41′ 44″ N	72′ 39″ W	1.7	2.43	3.48	4.07	5.14	5.58	5.38	5.04	4.13	2.91	1.81	1.42	3.59

- Thirty-four percent of the sun's energy is reflected back into space by snow and clouds. This reflective quality of a planet is called its *albedo*.
- Fourty-two percent of the energy goes to warm the land and water. The warmth of the earth is constantly being radiated into space, and the sun's energy replenishes this warmth.
- The water cycle—evaporation and precipitation—uses 23% of the solar energy.
- Winds and ocean currents use 1%.

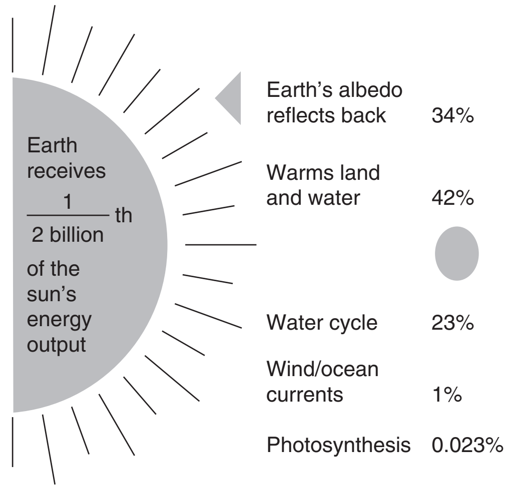


FIGURE 11-1 Percent of the sun's energy on the surface of the earth.

11.3.3 Flat-Plate Collector

A typical *flat-plate collector* (Fig. 11-2) is an insulated metal box with a glass or plastic cover, called the glazing, and a dark-colored absorber plate. Low-iron glass is a common glazing material for flat-plate as it gets high percentage of the total available solar energy. Simultaneously, only very little of the heat emitted by the absorber escapes the cover (greenhouse effect).

In addition, the transparent cover prevents wind and breezes from carrying the collected heat away (convection). Together with the frame, the cover protects the absorber from adverse weather conditions. Typical frame materials include aluminum and galvanized steel; sometimes fiberglass-reinforced plastic is used.

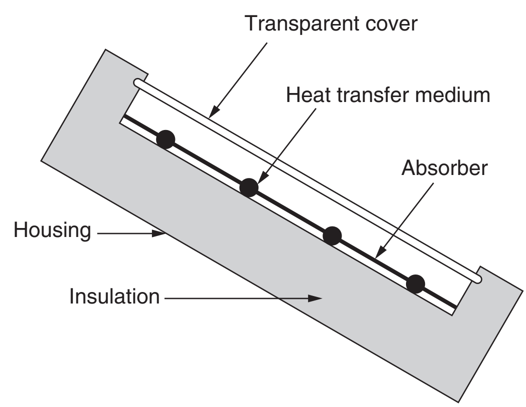


FIGURE 11-2 Elements of the basic flat-plate collector.

The insulation on the back of the absorber and on the sidewalls lessens the heat loss through conduction. Insulation is usually of polyurethane foam or mineral wool, though sometimes mineral fiber insulating materials like glass wool, rock wool, glass fiber, or fiberglass are used.

Flat-plate collectors fall into two basic categories: air and liquid. Both types can be either glazed or unglazed (according to the previously mentioned classification).

11.3.4 Collector Efficiency

The collector performance is evaluated in terms of collector efficiency, the ratio of the energy collected to that incident on the collector, usually expressed as a percentage. At a fixed rate of solar insulation, the collector efficiency of a given collector decreases with temperature difference between the collector and the surrounding air. Thus, there is a trade-off between temperature of collection and amount of energy collected. If high collection temperatures are desired, a larger amount of collector surface is needed than would be required to gather the same amount of energy at a lower collection temperature. Since the major cost of most solar energy systems is in the collectors, it is important to keep both the unit cost of the collectors and the total amount of collector surface as small as possible.

TABLE 11-3 Collection Efficiencies for Flat-Plate Collectors

Type	Day-long efficiency
Plastic tube type	25%
Bare plate	30%
Covered plate	35%
Suspended plate	40%
2-cover suspended plate	45%

Ordinarily, a surface that is a good absorber is also a good emitter. Collector efficiency can be improved by the use of a selective surface, one which has a high absorption for sunlight but a low radiation emittance. Selective surfaces are usually prepared by a plating on deposition process. At a fixed collector temperature, collector efficiency may be more than doubled over that of an ordinary surface when selective surfaces are utilized. Selective surfaces also permit the collection of energy at a higher than normal temperature without as large a decrease in collector efficiency as would occur for a nonselective absorber surface.

Table 11-3 presents collection efficiencies for flat-plate collectors.

11.3.5 Heating with Solar Energy

Today, we are able to harness to power of the sun in numerous ways, including using it for water heating, space heating, and space cooling in buildings. Many buildings are now designed to take full advantage of the sun's warmth, and incorporating solar heating in a building will begin to return on the investment immediately. Solar heating can be fully integrated into a building during the design phase or an existing building can be retrofitted to take advantage of solar heating. Solar heating can be used to provide hot water or heat the air in a building (space heating). Solar heating can be either passive, such as simply using large windows to let in more light and warmth, or active, where specially designed mechanical systems increase the heat gained from the sunlight.

Passive Solar Heating. Just as the name implies, passive solar heating allows the sun to do all the work. That is, there is no additional mechanical assistance. When referring to space heating, passive solar design takes advantage of the sun's warmth through such design features as large, south-facing windows and materials in the floors or walls that will absorb warmth during the day and release that warmth at night, when the heat is most needed because the south side of a building always receives the most sunlight. Passive solar water heating refers to a hot water system that is not aided by heat pumps. These systems will include a solar collector to heat the water and a storage tank to store the hot water.

Active Solar Heating. Active solar heating uses concepts similar to passive solar heating. However, active solar takes the power of the sun and amplifies it. Using specially designed mechanical systems, active solar heating can generate much more heat for space heating and hot water than passive solar alone. There are two basic types of active solar heating systems, depending on whether air or a liquid is heated in the solar collector. A liquid-based system heats water or an antifreeze solution in a "hydronic" collector, and an air-based system heats air in an "air" collector. Both of these systems collect and absorb solar radiation, then transfer the solar heat directly to the interior space or to a storage system; an auxiliary or backup system provides the additional heat.

The technical potential for residential applications of solar heating systems is 0.5 to 1.0 m^2 of solar collector/inhabitant. "Solar countries" such as Israel, Greece, and Cyprus already have high "solar water heating penetration" (Israel has about 0.95 m^2 per inhabitant), whereas some of the best IEA (International Energy Agency) countries, such as Greece and Austria, have a penetration of between 0.2 and 0.25 m^2 per inhabitant. The average solar penetration in IEA countries is roughly 0.04 m^2 per inhabitant; this would suggest that a strong growth of solar heating installations could be expected in the future. Driving forces for market development will be reduced costs and the desire to reduce greenhouse gas emissions. A number of studies carried out by the IEA and the European Commission in several countries reach a common conclusion: the market for solar water heaters is huge and—taken as a whole—is steadily growing, although the market growth will differ widely from country to country. Currently, the most important solar application is for residential water heating. Today, systems for hot water production in single-family houses are dominant; although, in the future, solar heating systems will be used in all types of housing. In countries with centralized heating systems, such as district heating, large-scale solar energy systems will feed heat to the distribution network. Such systems have been successfully demonstrated in Scandinavia and Germany. Swimming pool solar systems, common in some countries, also present a large market.

11.3.6 Solar Thermal-Conversion Plants

The thermal energy collected from a solar collector can be converted into work or mechanical energy by the use of a heat engine, which can then be used to generate electricity. The three types of thermodynamic cycles which seem to be practical with solar systems are the Rankine cycle, the Brayton cycle, and the Sterling cycle. The *Rankine cycle* is a vapor power cycle that is used with modifications in most large-sized electric generating plants. The *Brayton cycle* is a gas power cycle used as the basis of most gas-turbine power plants. The *Sterling cycle* is a high-efficiency cycle used as the basis for an external combustion gas engine with relatively low pollution and noise characteristics. A schematic flow diagram for a solar power plant operating on the Rankine cycle is shown in Fig. 11-3.

The limitations imposed by the second law of thermodynamics apply to any solar thermal power cycle. This limitation is best expressed in terms of the Carnot principle, which says that it is impossible to construct an engine that operates between two given heat reservoirs and which has a higher thermal efficiency than a Carnot engine operating between the same reservoirs. The *Carnot engine* is an idealized heat engine that is thermodynamically reversible and receives heat from a high-temperature reservoir (the source) and rejects heat to a lower-temperature reservoir (the sink). The useful work done per cycle is the difference between the heat added and the heat rejected during the cycle.

The thermal efficiency, the ratio of useful work done to the heat supplied, is expressed for the Carnot cycle in terms of the temperature of the reservoirs with which it is exchanging heat.

$$\eta = 1 - \frac{T_L}{T_H} \tag{11-5}$$

where η = thermal efficiency of the Carnot cycle
T_L = absolute temperature (degrees Celsius + 273.15°C) of sink
T_H = absolute temperature of source

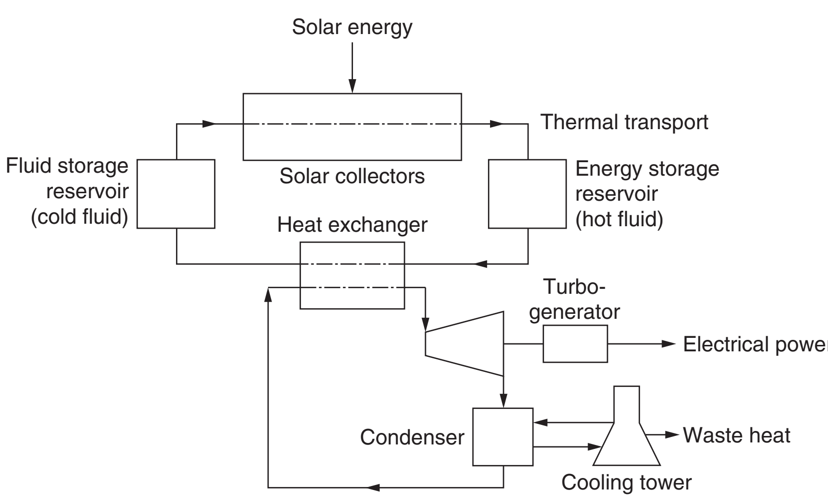


FIGURE 11-3 Schematic diagram of a solar power plant operating on the Rankine cycle.

For a solar energy system collecting heat at 121.1°C (250°F), the maximum thermal efficiency for any heat engine using that heat and rejecting heat to the atmosphere at 10°C (50°F) would be

$$\eta = 1 - \frac{273.15 + 10}{273.15 + 121.1} = 0.282 \qquad (11\text{-}6)$$

In other words, even an ideal heat engine would convert only 28% of the solar energy collected if the collector exit temperature were 121.1°C (250°F). A real engine would convert considerably less.

The White Paper of the European Commission for a community strategy and action plan on renewable energies of 1997 foresees at least 1 GWe of those systems implemented in Europe by the year 2010. This objective can be achieved by a scenario of a number of 25 to 30 commercial solar thermal power plants with 30 to 50 MWe unit size each and distributed along the south of Europe.

India, Egypt, Morocco, and Mexico have now applied within the GEF (Global Environment Facility) Operational Programme for about U.S. $50 million GEF grant for each project to cover their incremental costs of solar thermal power projects.

With positive experiences in construction and operation of the first European demonstration power plant projects being under development (50 MWe THESEUS on the Crete island in Greece; 10 MWe Planta Solar [PS 10] in Southern Spain), other projects are expected to follow. Until the year 2015, the market potential for solar thermal power plants is estimated at least with 7 GWe in Southern Europe, representing a CO_2 reduction potential of up to 12 million tons per year.

11.3.7 Concentrating Collectors

Flat-plate collectors have very poor collector efficiencies at temperatures above 93.3°C (200°F), and therefore are not suitable for use in solar thermal power plants. Collectors that concentrate the sun's rays can give higher collector efficiencies and much higher outlet temperatures than flat-plate collectors.

Concentrating collectors use mirrored surfaces to concentrate the sun's energy on an absorber called a *receiver*. The mirrored surface focuses sunlight collected over a large area onto a smaller absorber area to achieve high temperatures.

These collectors reach much higher temperatures than flat-plate collectors. However, concentrators can only focus on direct solar radiation, so their performance is poor on hazy or cloudy days. Concentrators are most practical in areas of high exposure to the sun's rays.

Concentrators are used mostly in commercial applications because they are expensive. Some residential solar energy systems use parabolic-trough concentrating systems. These installations can provide hot water, space heating, and water purification.

11.3.8 Central and Distributed Systems

Solar thermal power plants may be classified as central receiver systems or as distributed systems. In the central receiver system, solar energy is transferred optically from the individual collectors to a single receiver, for example, a boiler for a Rankine-cycle-type power plant. The most common approach for this type of plant is to locate the boiler at the top of a tall tower and to surround the tower with hundreds of mirrors which can reflect the sun's rays to the top of the tower. Systems have been proposed that would generate superheated steam at about 537.7°C (1,000°F) to reach thermal efficiencies comparable with conventional fossil-fueled power plants. The turbine generator would be located on the ground near the base of the tower.

In a distributed system, energy is transported from individual solar collectors by heated fluid flowing through pipes to a central boiler. The collectors are normally of the concentrating type such as parabolic, troughs, or paraboloidal dishes. Flat-plate collectors could be used, but the relatively low temperatures which they can produce lead to low thermal efficiencies. The resulting low overall plant efficiencies would require some rather large areas if flat-plate collectors were used. This is balanced to some degree by the lower cost per unit area of flat-plate collectors as compared with concentrating collectors.

Because of the intermittent nature of the sun, a solar power plant should be operated as an energy-displacement system, in connection with a conventional power system, which is employed whenever sufficient solar energy is not available because of clouds or night. Another approach is to have fossil fuel available to the solar system to furnish the required heat as needed. A third approach is to store thermal energy collected by the solar system for use during the periods of inadequate solar insulation. In a solar power plant this would involve the storage of a liquid at high temperature. A suitable storage medium should be low in cost and have high heat capacity, high temperature capability, and a lower vapor pressure. It also should be noncorrosive, nontoxic, and have a high thermal conductivity so that heat can be stored or removed at the desired rate without excess heat-transfer surface. If the fluid is to circulate through the solar collectors, it is also desirable that it does not freeze at the temperatures that might be encountered at night. Thermal insulation of the storage tanks is necessary, since storage will be at high temperatures and for reasonably long periods of time. Energy storage also could be accomplished by generating electricity and then using the electricity for pumped storage, electrolysis of hydrogen, charging of batteries, or flywheels.

11.3.9 Solar Energy Facts

In the United States. The position taken by the U.S. government in the current National Energy Review and the ability for the solar industry to structure long-term contracts will be key factors influencing the period over which solar energy takes to become economically self-sustaining. Proposals for a 10% federal tax credit up to $2,000 per solar system are under review.

Around 1.2 million solar thermal systems have been installed in the United States. Over 80% of these have been to residential users.

On June 10, 2001, Governor George Pataki of New York mandated that state facilities purchase at least 10% of their power needs from renewable sources by 2005, and 20% of those needs by 2010.

On May 16, 2001, the California Energy Commission increased the rebate on solar energy systems to the lesser of 50% off the purchase price or $4,500 per kilowatt (peak). Previous rebates based on levels as high as $3,000 per kilowatt (peak).

In Japan. For the fiscal year 2001, the Japanese solar rooftop program received applications for 114 MW of solar from 29,389 households.

Nearly 45% of the world's solar cell production is manufactured in Japan. Japan and the United States are the two biggest exporters of photovoltaic (PV) cells and modules.

Solar capacity is expected to increase nearly tenfold by 2010, which would then account for 30% of renewable energy supply. The national target is 5,000 MW of installed PV systems by the fiscal year 2010.

In Germany. The "Feed-in Law" in Germany permits customers to receive up to 45.7 eurocents per kilowatt hour for solar generated electricity. The program now calls for a total of 1,000 MW to be installed. By the end of 2001, the Kreditanstalt fur Wiederaufbau (KfW) Bank, which administers the 100,000 Roof Program in Germany, had approved loans for over 126 MW of PV systems. Bavaria tops the list of states in Germany with over 50 MW of systems approved.

The Feed-in Law fixes tariffs for approved renewable energy projects for a 20-year period from the plant commissioning and will apply incremental price cuts. Initial prices were set at 47.7 cents per kilowatt hour for solar energy, 8.6 cents per kilowatt hour for wind, from 9.6 to 8.2 cents per kilowatt hour for biomass, 8.4 to 6.7 cents per kilowatt hour for geothermal, and 7.2 to 6.3 cents per kilowatt hour for hydropower, waste, and sewage gas.

Some 20,000 solar electricity systems yielding an output of about 77 MW were installed in 2001, almost twice as many as the previous year. With these additions, the total solar electricity capacity in Germany is now estimated at over 170 MW.

According to current estimates, the German solar market reached a volume of about 1.5 billion marks in the year 2000. By the end of the decade, the market volume should increase to almost 7 billion marks.

In Australia. Renewable energy consumed as a percentage of the total energy consumed fell slightly between 1977 to 1978 and 1997 to 1998 (from 7% to 6% of primary energy consumed). Of this 0.9% was solar.

Worldwide. Solar energy accounts for less than 0.1% of the total global primary energy demand. Solar energy demand has grown at about 25% per annum over the past 15 years. (Hydrocarbon energy demand typically grows between 0% and 2% per annum.)

Japan, Germany, and the United States constituted 71% of the world market, unchanged on the previous year. In Japan and Germany, grid-connected applications accounted for over 95% of the market.

For the fiscal year 2002, the Japanese solar rooftop program received applications from 42,838 households.

Over 45% of the world's solar cell production is manufactured in Japan. Europe is second with 25% and the United States with 19%. Around 70% of the U.S. production is exported. However, this percentage will decline over the medium term.

Four companies account for over 50% of solar cell production: Sharp, Kyocera, BP Solar, and Shell Solar.

A residential solar energy system typically costs about $8 to $10 per watt. Where government incentive programs exist, together with lower prices secured through volume purchases, installed costs as low as $3 to $4 per watt or some 10 to 12 cents per kilowatt hour can be achieved. Without incentive programs, solar energy costs (in an average sunny climate) range between 22 and 40 cents per kilowatt hour for very large PV systems.

Japan overtook Switzerland in 2001 in terms of the proportion of solar cells installed per person in the country.

11.4 PHOTOVOLTAICS

BY AHMED ZOBAA

Solar electric or PV systems convert some of the energy in sunlight directly into electricity. PV cells are made primarily of silicon, the second most abundant element in the earth's crust and the same semiconductor material used for computers. When the silicon is combined with one or more other materials, it exhibits unique electrical properties in the presence of sunlight. Electrons are excited by the light and move through the silicon. This is known as the *photovoltaic effect* and results in dc electricity. PV modules have no moving parts, are virtually maintenance-free, and have a working life of 20 to 30 years.

There are three basic categories of PV systems with several types in each category. Crystalline silicon flat-plate collectors are the most developed and prevalent type in use today. These include single-crystal silicon and polycrystalline silicon, which are either grown or cast from molten silicon and later sliced into their cell size. They are then assembled onto a flat surface; no lenses are used. Thin-film systems are inherently cheaper to produce than crystalline silicon but are not as efficient. They are produced by depositing a thin layer of PV material to a substrate such as glass or metal. This group includes amorphous silicon, like the kind found in calculators and watches. Concentrators use much less of a specialized PV material and employ a lens or reflectors to concentrate sunlight on the PV cell and increase its output. They can be produced more cheaply than either of the other type due to the reduced amount of expensive PV material. However, they can use only direct sun, so they must track the sun precisely and do not work when it is cloudy.

11.4.1 Photovoltaic System Terms

PV system terms progress from small to large as follows:

- PV cells, the smallest unit of a PV system, are wired together to form modules.
- *Modules* are usually a sealed or encapsulated unit of convenient size for handling.
- Modules are wired together to form panels.
- Groups of panels form arrays.

- A number of arrays form an array field.
- The total system includes the arrays and any other equipment, such as charge controllers, storage (batteries), tracking, and monitoring equipment, collectively called *balance of system (BOS) components*.

11.4.2 History of Photovoltaics

The history of photovoltaics dates back to 1839, and major developments evolved as follows:

- In 1839, Edmund Becquerel, a French physicist, observed the photovoltaic effect.
- In the 1880s, selenium PV cells were built that converted light in the visible spectrum into electricity and were 1% to 2% efficient. Light sensors for cameras are still made from selenium today.
- In the early 1950s, the Czochralski meter was developed for producing highly pure crystalline silicon.
- In 1954, Bell Telephone Laboratories produced a silicon PV cell with a 4 % efficiency and later achieved 11% efficiency.
- In 1958, the U.S. Vanguard space satellite used a small (less than 1-W) array to power its radio.
- The space program has played an important role in the development of photovoltaic ever since.
- During the 1973 to 1974 oil embargo, the U.S. Department of Energy funded the Federal Photovoltaic Utilization Program, resulting in the installation and testing of over 3,100 PV systems, many of which are still in operation today.
- The 1970s through the 1990s have seen a relative disinterest in solar power, with majority ownership of many U.S. PV manufacturers being transferred to German and Japanese interests.
- The Gulf War of 1990 again sparked America's interest in non-fossil fuel energy alternatives.

11.4.3 The PV Power Market

PV systems traditionally have been economical in remote applications. Most common examples include wireless and cellular communications systems, off-grid homes, recreational vehicles and boats, power for offshore oil rigs, and highway sign lighting and call boxes. Water pumping, vaccine refrigeration, and water purification have all been important roles for photovoltaics in developing countries.

Market forces seem to have a hold of the PV market, since sales in 1995 rose 20%. Most U.S. manufacturers are increasing production significantly, and costs are expected to fall with the new volumes.

Current estimates of worldwide production of solar photovoltaic cells and modules for 1998 are about 120 MW, up steadily and dramatically from only 40 MW in 1990. Worldwide sales have been increasing at an average rate of about 15% every year during the last decade, although that growth rate has been slower in some markets and regions but faster in others. We believe that there is a realistic possibility for the market to continue to grow at about a 15% rate into the next decade. At this rate, the world production capacity would be 1,000 MW by 2010, and photovoltaics could be a $5 billion industry. These are realistic benchmarks, and show the solar business to be a very exciting market opportunity in the near term.

Developing countries today are the largest and fastest growing segment of the PV market. For the 2,000 million people in the developing world who currently have no access to basic electrical services, PV presents the opportunity for a giant leap forward and a much needed improvement in living standards. For the PV services industry, the developing world represents an enormous new business opportunity.

Photovoltaic prices are continuing a downward trend as manufacturers increase production. Cost decreases, combined with national and state incentive and subsidy programs, and renewable energy capacity mandates, have led to the emergence of a steadily growing market for bulk photovoltaic installations (greater than 100 kW capacity). Merchant PV plants are being built in places where favorable feed-in tariffs make projects profitable. Bulk installations significantly decrease the installed cost of PV power at individual sites, while simultaneously driving market demand for PV equipment. The major players and PV technologies in each market are identified along with the mounting modes (roofing tiles, weather skins, carport shading, window walls, and so on) that are prominent. The potential impact of mass production of promising emerging PV technologies is examined. Market forecasts are provided for capacity, new projects, and annual revenue for the 2002 to 2008 time frame. The forecasts cover national, regional, and world markets for both single-site bulk PV installations and bulk purchases for national village power supplies and large residential projects. Project revenue is growing at over 40% per year, and will likely reach $1 billion annually by 2010.

International Activity. U.S. PV exports increased in capacity from 14.814 million ft^2 in 1993 to 17.714 million ft^2 in 1994, an increase of 19.6%. U.S. PV imports increased from 1.767 million ft^2 in 1993 to 1.98 million ft^2 in 1994, an increase of 12%. U.S. module production is leading world growth as well. In 1993, the United States produced 21 MW of PV, of which 14.8 MW was exported. By 1997, global demand led to a record breaking PV production year with a 42% leap in worldwide production. The United States produced 46.4 MW with $175 million in sales and exported 33.8 MW (73% of production) overseas. India is boosting production and becoming a major world producer of PV modules. The Indian government plans to power 100,000 villages with renewable energy, primarily PV modules, and install solar-powered telephones in each of the nation's 500,000 villages. Mexico planed to electrify 60,000 villages using photovoltaics by the year 2000. Hospital Bulape (serving 50,000 outpatients per year in Zaire) and several other major hospitals in Zaire depend totally on solar power for everything from x-ray equipment to air conditioning. In Morocco, solar panels are sold in bazaars and open markets, next to carpets and tinware. In San Buenaventura, Guatemala, a local utility has installed PV panels on 42 of the community's 86 homes at one-third the cost of extending power lines into the village. Malta-Solar Power, Ltd. has begun construction of a new PV plant with a maximum production capacity of 3 MW per year. At full production, this plant will be able to produce 40% of all 1995's module production capacity for all developing countries. South African companies are building a PV manufacturing plant near Alexandra township that will serve to electrify 10,000 homes, 600 clinics, and 1,000 schools with solar power. Kenya has electrified 20,000 homes using photovoltaics in the last few years, compared with 17,000 new homes that were hooked up to the central power grid. Siemens Solar in 1995 sold just over 40% of its output in North and South America, nearly 40% to Europe and Africa, and "just under" 20% went to Asia. 5SI President Gernot Oswald expects the biggest growth in the next few years to come from Asia. There are over 500,000 homes using PV today in villages around the world for electricity. In Kenya, more rural households receive electricity from PV than from the conventional power grid. The single largest market sector for PV is village power at about 45% of worldwide sales. This is mostly comprised of small home lighting systems and water pumping. Remote industrial applications such as communications are the second largest market segment.

11.4.4 Global PV Market

The fast growing world market for PV greatly reflects the growing rural electrification demand of less developed countries around the world. The global PV market has grown at an average rate of 16% per year over the decade with village power driving demand. The total worldwide PV production in 1980 was only 6.5 MW, and by 1997 this had increased to 126.7 MW.

For many applications, especially remote site and small power applications, PV power is the most cost-effective option available, not to mention its environmental benefits. New PV modules generally retail for about $5 per peak watt, depending on quantities purchased. Batteries, inverters, and other balance of system components can raise the overall price of a PV system to over $10 to $15 per installed watt. PV modules on the market today are guaranteed by manufacturers from 10 to 20 years, while many of these should provide over 30 years of useful life. It is important when designing PV systems to be realistic and flexible, and not to overdesign the system or overestimate energy requirements (e.g., overestimating water-pumping requirements) so as not to have to spend more money than needed. PV conversion efficiencies and manufacturing processes will continue to improve, causing prices to gradually decrease.

PV conversion efficiencies have increased with commercially available modules that are from 12% to 17% efficient, and research laboratory cells demonstrate efficiencies above 34%. A well-designed PV system will operate unattended and requires minimum periodic maintenance, which can result in significant labor savings. PV modules on the market today are guaranteed by the manufacturer from 10 to 25 years and should last well over 30 years. PV conversion efficiencies and manufacturing processes will continue to improve, causing prices to gradually decrease; however, no dramatic overnight price breakthroughs are expected.

11.4.5 Common Photovoltaic Applications

PV is best suited for remote site applications that have small to moderate power requirements, or small power consuming applications even where the grid is in existence. A few power companies are also promoting limited grid-connected PV systems, but the large market for this technology is for stand-alone (off-grid) applications. Some common PV applications are as follows:

Water Pumping. Pumping water is one of the most competitive arenas for PV power since it is simple, reliable, and requires almost no maintenance. Agricultural watering needs are usually greatest during sunnier periods when more water can be pumped with a solar system. PV-powered pumping systems are excellent for small to medium scale pumping needs (e.g., livestock tanks) and rarely exceed applications requiring more than a 2 hp motor. There are thousands of agricultural PV water pumping systems in the field today throughout Texas. PV pumping systems' main advantages are that no fuel is required and little maintenance is needed.

A PV-powered water pumping system is similar to any other pumping system, only the power source is solar energy; PV pumping systems have, as a minimum, a PV array, a motor, and a pump. PV water pumping arrays are fixed mounted or sometimes placed on passive trackers (which use no motors) to increase pumping time and volume. AC and dc motors with centrifugal or displacement pumps are used with PV pumping systems. The most inexpensive PV pumpers cost less than $1,500, while the large systems can run over $20,000. Most PV water pumpers rarely exceed 2 hp in size. Well installed quality PV water pumping systems can provide over 20 years of reliable and continuous service.

Gate Openers. Commercially available PV-powered electric gate openers use wireless remote controls that start a motorized actuator that releases a gate latch, opens the gate, and closes the gate behind the vehicle. Gates are designed to stop if resistance is met as a safety mechanism. Units are available that can be used on gates up to 16 ft wide and weighing up to 250 lb. Batteries are charged by small PV modules of only a few watts. Digital keypads are available to allow access with an entry code for persons without a transmitter. Solar-powered gate-opening assemblies with a PV module and transmitter sell for about $700.

Electric Fences. P-power can be used to electrify fences for livestock and animals. Commercially available packaged units have maintenance free 6 or 12-V sealed gel cell batteries (never need to add water) for day and night operation. These units deliver safe (non-burning) power spikes (shocks) typically in the 8,000 to 12,000 V range. Commercial units are UL (Underwriters Laboratories) rated and can effectively electrify about 25 to 30 miles of fencing. Commercially packaged units are available from about $150 to $300, depending on voltage and other features.

Water Tank Deicers. For the north plains of Texas in the winter, PV power can be used to melt ice for livestock tanks, which frees a rancher from going out to the tank with an ax to break the surface ice so the cows can drink the water. The PV module provides power to a small compressor on the tank bottom that generates air bubbles underwater, which rise to the surface of the tank. This movement of the water with the air bubbles melts the tank's surface ice. Commercially available units are recommended for tanks 10 ft in diameter or greater, and can also be used with ponds. Performance is best for tanks that are sheltered, bermed, or insulated. Installation is not recommended for small, unsheltered tanks in extremely cold and windy sites. Approximate cost for a complete owner-installed system, including a PV module, compressor, and mounting pole, is about $450.

Commercial Lighting. PV-powered lighting systems are reliable and a low-cost alternative widely used throughout the United States. Security, billboard sign, area, and outdoor lighting are all viable applications for PV. It's often cheaper to put in a PV lighting system as opposed to installing a grid lighting system that requires a new transformer, trenching across parking lots, etc. Most stand-alone PV lighting systems operate at 12 or 24 V dc. Efficient fluorescent or sodium lamps are recommended for their high efficiency of lumens per watt. Batteries are required for PV lighting systems. Deep cycle batteries specifically designed for PV applications should be used for energy storage for lighting systems. Batteries should be located in protective enclosures, and manufacturer's installation and maintenance instructions should be followed. Batteries should be regulated with a quality charge controller. Lighting system prices vary depending on the size; average systems cost from $600 to $1,500.

Residential Power. Over 500,000 homes worldwide use PV power as their only source of electricity. In Texas, a residence located more than a mile from the electric grid can install a PV system more inexpensively than extending the electric grid. A Texas residence opting to go solar requires about a 2 kW PV array to meet its energy needs, at a cost of about $15,000. The first rule with PV is always energy efficiency. A PV system can provide enough power for an energy-efficient refrigerator, lights, television, stereo, and other common household appliances.

A great number of PV installations for homes have taken place in Mexico. The experience of PV electrification varies widely across Mexico and is demonstrative of the potential pitfalls of haphazard installations. Over 40,000 PV home lighting systems have been installed in Mexico, mostly through government programs (Foster 1998). However, nearly half of these systems are not functioning today, mostly due to poor balance of systems hardware (i.e., the PV modules work fine), where improper batteries and poor quality charge controllers are used. It is important for any PV user to use quality equipment and install PV systems in accordance with local electric codes. This greatly reduces the potential for future problems.

Evaporative Cooling. PV-powered packaged evaporative cooling units are commercially available and take advantage of the natural relation that when maximum cooling is required is when maximum solar energy is available. These units are most appropriate for comfort cooling in the dry climate of West Texas where performance is best. Direct evaporative coolers save 70% of the energy over refrigerated units. Battery storage is obviously required if cooler operation is desired at night. Array size would vary with the power requirements of the cooler motor. A linear current booster (LCB) is useful between the PV modules and the cooler's dc motor if the cooler is coupled directly to the PV array. Packaged PV evaporative cooling systems for residences generally run from $500 to $1,500, depending on size.

Telecommunications. This was one of the early important markets for PV technologies, and continues to be an important market. Isolated mountaintops and other rural areas are ideal for stand-alone PV systems where maintenance and power accessibility make PV the ideal technology. These are often large systems, sometimes placed in hybrid applications with propane or other type of generators.

Consumer Electronics. Consumer electronics that have low power requirements are one of the most common uses for PV technologies today. Solar-powered watches, calculators, and cameras are all everyday applications for PV technologies. Typically, these applications use amorphous PV technologies that work well even in artificial light environments such as offices and classrooms.

11.4.6 Glossary of Solar and Photovoltaic Terms

Cell Efficiency. The ratio of the electrical energy produced by a PV cell (under full sun conditions or 1 kW/m^2) to the energy from sunlight falling on the cell.

Charge Controller. A component that controls the flow of current to and from the battery subsystem to protect the batteries from overcharge and overdischarge. The charge controller also may monitor system performance and provide system protection.

Diffuse Radiation. Sunlight received indirectly as a result of scattering due to clouds, fog, haze, dust, or other substances in the atmosphere.

Direct Radiation. Light that has traveled in a straight path from the sun (also referred to as beam radiation). An object in the path of direct radiation casts a shadow on a clear day.

Flat-Plate Array. A photovoltaic array in which the incident solar radiation strikes a flat surface and no concentration of sunlight is involved.

Fresnel Lens. A concentrating lens positioned above and concave to a PV material to concentrate light on the material.

Grid-Connected. Referring to an energy-producing system connected to the utility transmission grid (also called utility interactive).

Hybrid System. A power system consisting of two or more power-generating subsystems (e.g., the combination of a wind turbine and a PV system).

Insolation. The amount of sunlight reaching an area usually expressed in watts per day.

Load. Electrical power being consumed at any given moment. The load that an electrical generating system supplies varies greatly with time of day and to some extend season of year. Also in an electrical circuit, the load is any device or appliance that uses power.

Parallel-Connected. Referring to a method of connection in which positive terminals are connected together and negative terminals are connected together. Current output adds and voltage remains the same. (See also series-connected.)

Photovoltaic Cell. The semiconductor device that converts light into dc electricity. The building block of PV modules.

Series-Connected. Referring to a method of connection in which the positive terminal of one device is connected to the negative terminal of another. The voltages add and the current is limited to the least of any device in the string. (See also parallel-connected.)

Solar Constant. The rate at which energy is received from the sun just outside the earth's atmosphere on a surface perpendicular to the sun's rays. Approximately equal to 1.36 kW/m^2.

Thick Cells. Conventional cells, such as crystalline silicon cells, which are typically from 4 to 17 mil thick. In contrast thin-film cells are several micrometers thick.

Thin-Film Cells. PV cells made from a number of layers of photosensitive materials. These layers are typically applied using a chemical vapor deposition process in the presence of an electric field.

Voltage Regulator. A device that controls the operating voltage of a PV array.

11.5 WIND POWER

BY PALMER CARLIN

11.5.1 Introduction

In order to present the current status of wind energy technology concisely but still include important related issues, the following material is divided into four sections: the first section provides a quick survey of contemporary activity in wind energy. The second section presents the basic physics and technology of wind energy conversion together with a description of turbine features. The third section includes issues that arise in the application of wind machines, and the fourth and final section contains conclusions and references.

11.5.2 Contemporary Activity in the Wind Energy Field

With the onset of the energy crisis in 1976, interest in renewable energy sources suddenly intensified. As a result, both private and government agencies became involved. As an example of private activity, the American Wind Energy Association (AWEA) was founded with provision for both corporate and individual memberships. Typical corporate members are electric utilities and wind turbine manufacturers, while typical individual members are wind consultants, university personnel, wind turbine owners, and private citizens. A directory of its membership is available on the Internet.[1]

The U. S. Department of Energy (DOE), working through the National Wind Technology Center (NWTC) of the National Renewable Energy Laboratory (NREL) in Golden, Colo., supports a comprehensive wind energy program. Sandia National Laboratories (SNL) in Albuquerque, New Mexico, participates in the national program as well. One part of the program offers cost-shared contracts to private industry on a competitive basis for the development of advanced wind technology. Other parts of the program support wind research within the government laboratories, and the testing facilities at the NWTC can be made available for qualifying private companies.

In anticipation of the rising global importance of wind energy, in 1977 the International Energy Agency launched the Implementing Agreement for Co-operation in the Research and Development of Wind Turbine Systems.[2] Since then, 19 nations have joined this group by becoming members of its executive committee, which meets semiannually to exchange information about wind energy developments. The location of the meeting rotates among member nations, and the hosting nation includes visits to its wind installations. As wind technology progresses and new challenges arise, the committee creates suborganizations called *annexes* to which member countries contribute expert delegates. The operating agent of each annex reports to the executive committee at its semiannual meetings until the annex completes its task. In their 2004 Annual Report, the executive committee stated that the world's wind generation capacity exceeded 47 GW and was expected to continue growing at a rate of 28% per year.

11.5.3 Wind Turbine Analysis and Description

Wind Turbine Power Calculation. Good engineering models for predicting wind turbine performance can be obtained from some simple assumptions plus a few equations from elementary physics. By considering energy, momentum, and mass conservation laws, aerodynamicists in the mid-nineteenth century established the "axial momentum" theory for analysis of airplane propellers. One of the results is the expression for the flux of kinetic energy of the wind or any other fluid through an area normal to that fluid flow, which is

$$P_w = {}^1/_2\, \rho\, A\, V_w^{\,3} \tag{11-7}$$

where P_w is total wind power in watts, ρ is air density in kilograms per cubic meter, A is a mathematical area in square meters normal to the wind, and V_w is the wind speed passing through that area in meters per second. The cubed velocity is plausible because each air mass element carries kinetic energy proportional to velocity squared, and the rate of passage of each element being directly proportional to velocity yields the third velocity factor.

Because the extractable energy of the wind exists only as macroscopic kinetic energy, no practical machine can remove all of that energy, as there is no way to dispose of the resulting stationary air. Two aeronautical engineers independently examined this limitation in the 1920s. Lanchester[3] and Betz[4] separately used the ideal model in Fig. 11-4. It assumes that the perfectly inviscid fluid passes through an "actuator disc," which represents a wind machine rotor. An *actuator disc* is a mathematical convenience commonly used in wind turbine (as well as aircraft propeller) analysis. This porous disc extracts

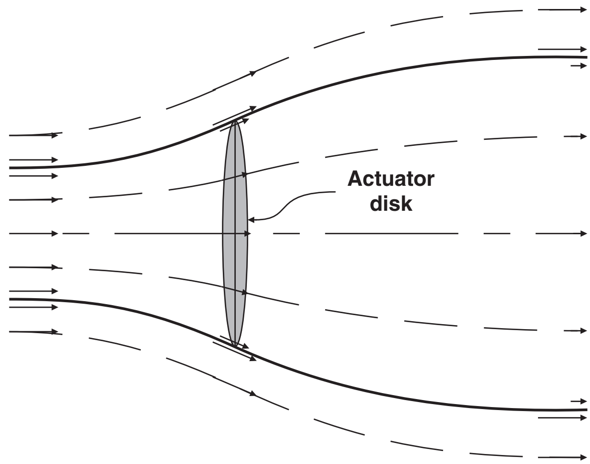


FIGURE 11-4 Idealized inviscid fluid flows through an actuator disc.

mechanical energy from the flow by causing a pressure drop. On the upstream side, the pressure has been raised above atmospheric by the slowing airstream, while on the downstream side pressure is lower. Ambient atmospheric pressure in the flow will be recovered downstream by further slowing. By extending axial momentum theory to this model, they showed that the maximum fraction of the above wind power that can be captured is $^{16}/_{27}$ or about 59%. This is known as the *Lanchester/Betz Limit* or more commonly the *Betz Limit*. Mathematical results derived using this model are assumed to apply as an upper limit to any device that extracts kinetic energy from a free fluid stream.

To account for this fundamental limit as well as additional imperfections introduced by practical wind turbines, wind engineers define the dimensionless power coefficient, C_p, for a wind turbine as follows

$$P = {}^1/_2\rho A C_p V_w^{\ 3} \qquad \text{watts} \tag{11-8}$$

where P is now the useful extracted power. One can see that this coefficient for a particular machine is the fraction of the total wind power that the wind machine can reclaim in a wind of V_w. Depending on the context this power may or may not include the losses in the power train of the turbine. Because of the Betz limit this fraction can never exceed 0.59. Measured values of power coefficient for real turbines range from 0.25 to more than 0.45.

There are a few important additional observations to be made in Fig. 11-4. First, note that the wind speed in the tube of air that will impinge on the disc (bounded by solid lines) begins to slow before it arrives there and it continues to slow after it leaves. As this fluid is assumed ideal, it has no viscosity, so the surrounding air slides past the subject tube and neither speed is affected by the other.

Second, for the mass flow to remain constant at any cross section, the declining fluid speed requires the tube cross section to expand. The velocity reduction and, therefore, the change in flow cross section proceeding from left to right, depends on the amount of energy being removed from the tube. Note that these results are independent of the type of device that is extracting energy from the flow. In the limit of no removal, the fluid tube cross section remains constant and equal to that of the actuator disc, and there would be no reduction of wind speed.

Finally, the velocity vectors shown represent the optimal case of the Betz limit. For this case, the impinging wind speed has been reduced to two-third of its original value as it reaches the turbine disc, and ultimately declines to one-third of the original value downstream. In this ideal case, the velocity reduction for this tube of fluid is permanent as mentioned above. In real cases of viscous flow in the atmosphere, this wake is dissipated and reenergized downstream by the surrounding unretarded wind. This wake process is very important to the design of a wind power plant with many successive rows of turbines. Turbines positioned farther into the center of a wind power plant are subjected to much greater turbulent buffeting that results in blade fatigue and up to 15% loss of annual energy compared to the most upwind turbines.

Obviously, current engineering analyses that account for the rotating blades and the effect of material fatigue are much more refined. For an excellent account of these approaches, consult Ref. 5.

11.5.4 Wind Turbine Classes

Based on the method of extracting energy from the airstream, two fundamental classes of wind machines can be defined. "Lift type" machines use airfoils that create lift at right angles to the airflow like an airplane wing. On the other hand, if a machine extracts energy by depending on the wind to push a specially shaped object directly downwind, it is a member of the other and much older class of wind machines called "drag machines."

In the lift class of machines, it is convenient to divide turbines into two main classes based on the orientation of their axis of rotation. The rotor axis of the horizontal axis wind turbine (HAWT) is parallel to the flow of the wind, while the rotor axis of the vertical axis wind turbine (VAWT) is transverse to the wind. At this writing, the latter constitutes less than 3% of commercial wind machine installations.

Horizontal Axis Wind Turbine. Usually, the designer of a new wind machine will select the power rating and the axis orientation first. For a HAWT at least four other important choices follow:

Rotor Orientation. In an upwind HAWT, the wind passes through the rotor before passing the tower, so that the blades are never shielded from the direct wind. However, special mechanical drives with control systems and direction sensors must be provided to move (yaw) and maintain the rotor in this upwind position. Conversely, a downwind machine can depend on the natural tendency of the wind to blow the rotor to the opposite side of the tower, and thus preclude the need for the special mechanical drive and sensors. The cost, however, is that each blade must pass through the turbulent wind region behind the tower, and thereby experience mechanical impulses that can shorten the rotor's fatigue life and create additional noise.

Blade Count. Although multiple-bladed turbines exist, economic factors at least for the present, have determined that either two- or three-bladed rotors are the most practical. It should be mentioned that the moment of inertia of a three-bladed rotor about its yaw axis is independent of the rotor's position, while this quantity varies from a maximum to a minimum as a two-bladed rotor rotates from horizontal to vertical orientation. During high yawing rates, this means that the gyroscopic moment on the main shaft is fluctuating wildly for a two-bladed rotor (See Hub Type).

Hub Type. The simplest and most obvious case of blade attachment is the rigid hub in which there is no relative motion of the rotor blades with respect to the supporting shaft. This is the usual case for smaller three-bladed rotors and is often used for two blades. Alternatively, suppose the wind machine rotor uses only two blades built as a long rigid beam, and suppose these blades are attached to the wind machine's main drive shaft by a single pin transverse to both this shaft and the blades. If this pin acts as a hinge allowing a small amount of movement, then this is a "teetering rotor." This feature eliminates bending at the end of the main rotor shaft in a turbulent wind and allows the plane of rotation of the blades to be somewhat misaligned with the main rotor shaft of the turbine for a few moments if necessary.

Aerodynamic Control. In considering aerodynamic control, we must distinguish between the control of power flow and the need to stop rotation in case of equipment failure or excessive wind. Although the speed of a wind turbine is normally controlled by the load torque of its electric generator, excess wind, loss of utility power, breakdown of the generator, or loss of the mechanical transmission would allow the turbine rotor to spin out of control. For this reason, all wind machines need some form of aerodynamic control.

The two most common hardware approaches to meeting this problem are (1) stall control and (2) pitchable blades. If stall control is chosen, the turbine will be designed such that when the wind exceeds the speed that fully loads the generator, the resulting large angle of attack of the wind on the blade airfoil causes the blades to stall thereby reducing mechanical power to the generator shaft.

In this case, the wind machine rotor blades are mounted permanently at a fixed pitch angle. The appeal of this method is its simplicity and consequent lower cost.

Alternatively, the existence of pitch control on propeller driven aircraft suggests full-span pitch control for wind machines, in which each blade can be rotated around its own longitudinal axis. When the rotation is sufficient to present the leading edge directly into the wind (i.e., approximately 90°), the average wind torque on the rotor will go to zero. Thus, positive control is achieved at the expense of rotating blade root attachments and complicated mechanical hub linkages, but enjoys the advantage (over stall control) of being able to reduce wind torque to zero on average.

Many stall-controlled turbines have additional aerodynamic control for emergencies in the form of tip flaps. These are flat plates mounted at the tip of the wind machine blades and normal to the blade's longitudinal axis. During normal machine operation they are in the stowed position where they present their edge to the slipstream and consequently offer low drag resistance. In case of incipient loss of control, these tip flaps can be deployed either by centrifugal force or by electrical trigger. They then pivot outward, presenting their entire surface broadside to the slipstream. This tangential drag force dramatically reduces rotor speed, though it does not stop the rotor completely. A mechanical brake is provided for that purpose.

Vertical axis wind turbine. The Darrieus or vertical axis wind turbine, which is sometimes called the "eggbeater" type machine, has several important advantages. It will operate with the wind approaching from any direction, the generator can be mounted in the base of the machine, and the primary mechanical loads on the blades are tension. Measured power coefficients are similar to those for HAWT machines. Some disadvantages are that aerodynamic torque only occurs as the blade moves across the wind, which results in torque pulsation, and the turbine is not self-starting. However, the primary problem is that the long slender blades are subject to many different modes of vibration with the resulting loss of fatigue life.

11.5.5 Wind Turbine Performance

Probably the most important quantitative information about a wind turbine is its power curve. A typical example is shown in Fig. 11-5 in which the rotor is assumed to be turning at a constant speed. The ordinate shows the power that a machine will produce if it is being driven by the steady wind shown on the abscissa. This example machine reaches its rated speed and power at about 17 m/s, and its furling or cutout wind speed at 23 m/s at which point its control system will automatically shut it off.

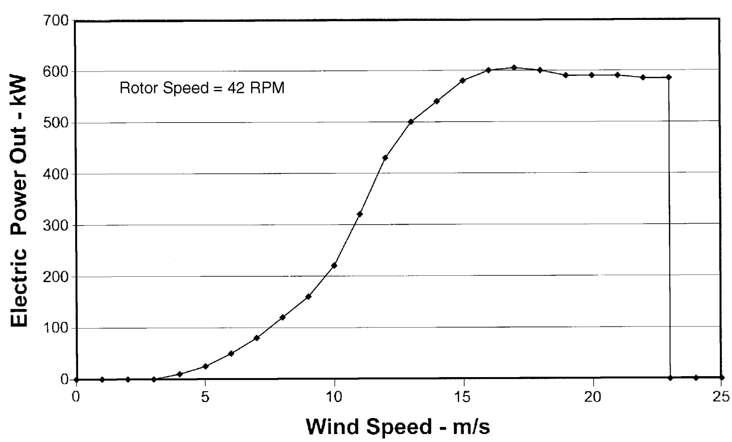


FIGURE 11-5 A typical power curve for a constant speed turbine (rotor speed = 42 rpm).

Although the previous power expression shows that power should be proportional to the cube of the wind speed, typical wind machine power rises more slowly and levels off at higher wind speed. This is because the power coefficient, C_p, is not a constant. To account for the effect of rotor speed on the power coefficient, it is useful to define a dimensionless quantity called "tip-speed ratio" sometimes abbreviated TSR or lambda (λ).

$$\lambda = R\omega/V_w \tag{11-9}$$

where R is the radius of the circle swept by the rotor, ω is the rotor angular speed in radians per second, and V_w is the wind speed. Note that the TSR can be thought of as the linear speed of a rotor blade tip measured in units of the existing wind speed. Thus, for a TSR of 7, the blade tip is traveling 7 times faster than the wind at that instant.

For a wide range of rotor and wind speeds it can be shown that the power coefficient is a function of TSR as is seen in Fig. 11-6. It is therefore possible, by knowing rotor speed and wind speed, to calculate the power produced by the given wind turbine using the expression

$$P = {}^1/_2\rho A C_p(\lambda) V_w^{\,3} \tag{11-10}$$

where C_p is obtained from Fig. 11-6. By finding the TSR for maximum C_p, we can calculate the rotor speed for a given wind speed that will give best performance for this machine. It will therefore mark the wind speed neighborhood in which we should strive to operate the machine.

This is discussed further in Sec. 11.5.8.

Figure 11-6 also shows that the power coefficient curve drops to zero for high TSRs. This can be interpreted as the TSR that would be obtained if no load were placed on the wind machine rotor, and the TSR at that point allows us to calculate the "runaway" rotor speed for any given wind. Although this speed is finite, most large machines are not designed to operate at such high speeds and would be damaged or destroyed if they were allowed to do so. On the other hand, for TSRs less than about 3, the power coefficient can be seen to be low and usually not well-known, because most fixed-pitch turbine airfoils will be stalled at this angle of attack.

While the information in Fig. 11-6 is sufficient to approximate the performance of fixed-pitch, constant-speed machines, if blade pitch is available as another control variable, then more information

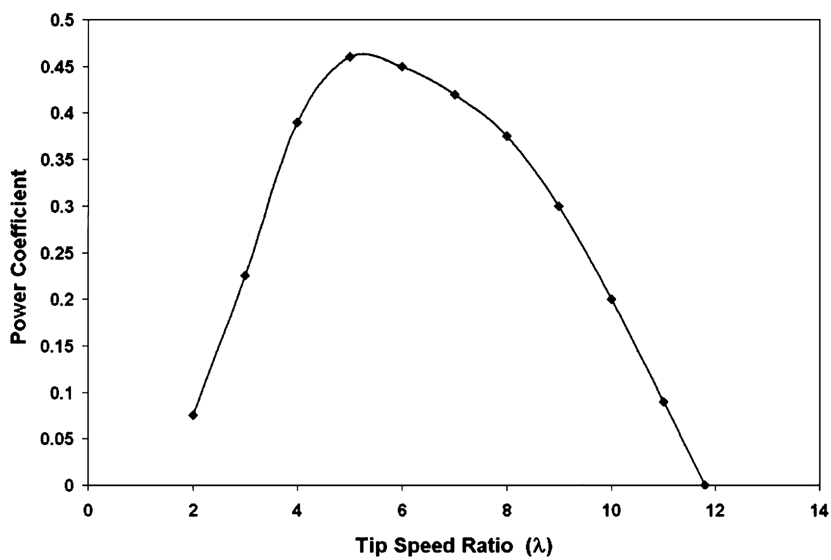


FIGURE 11-6 Example of a power coefficient vs. tip-speed ratio curve.

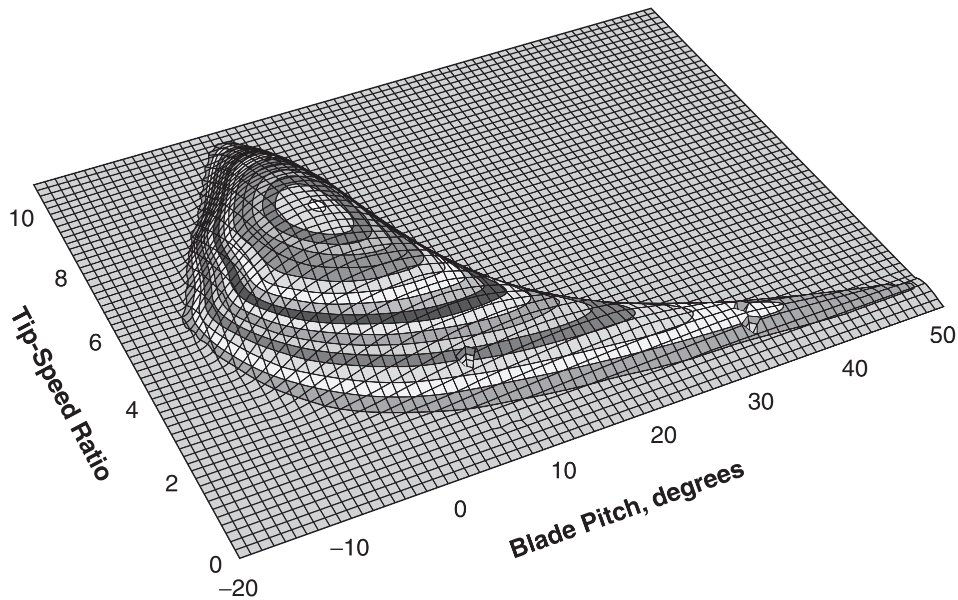


FIGURE 11-7 Power coefficient surface of a typical wind turbine.

is needed. Recall that blade-pitch angle is the angle between a blade chord line and the plane of rotation of the wind machine rotor, and for most high-speed turbines, it is in the range of 0° to 6°.

To account for this new variable, we must resort to a mathematical surface in three-dimensional space. In Fig. 11-7, the power coefficient is shown as a function of both TSR and blade-pitch angle. The fact that this surface has a single C_p maximum near zero degrees pitch and a TSR of 6 shows that maximum performance for this turbine can be had at really only one blade-pitch angle and at one TSR. Thus, ideal performance will occur only if the rotor blade tips are moving about 6 times faster than the existing wind speed and the blades are pitched to approximately zero. Of course, if a wind machine's electrical or mechanical capacities are being exceeded, pitch can be used to reduce power.

Other aerodynamic control techniques that have been tested include pitchable blade tips, ailerons, spoilers, and generator torque.

All of the preceding material applies only to devices that use airfoils that extract energy through the use of aerodynamic lift. The older wind machines that use the drag phenomenon exist in many forms, such as the ubiquitous cup anemometer. Although the latter is useful for accurate wind speed measurement, it is not efficient in the collection of raw energy. Machines using this technology are usually robust and can be fabricated using unsophisticated equipment. However, per unit of active collection area, they tend to require much more material than a lifting machine. But more important is that the maximum theoretical power coefficient for this technology ($Cp = {}^{4}/_{27}$) is exactly one-fourth that of the lifting class of wind machine ($Cp = {}^{16}/_{27}$). This is the reason that drag machines are never found in commercial wind power plants.

The purpose of this hardware section was limited to a discussion of wind machine energy conversion performance. The complete detailed design of a commercial wind machine would also include an extensive structural loads and fatigue analysis.

11.5.6 The Wind Resource

As indicated in the previous discussion, the wind turbine's energy source, the wind, must be well-known to accurately site turbines and predict performance. The Pacific Northwest National Laboratory (PNNL), at the request of DOE, began studying the wind resource in the United States and its territories in 1979 to identify practical locations for wind energy exploitation. Today, wind resource activities take place at NREL. In the wind maps developed by PNNL and NREL, the results are presented in terms of wind classes. Class 1 is least energetic while class 7 is the most energetic. An annual average wind resource map for the United States is shown in Fig. 11-8. In the atlas developed by PNNL[6], wind resource data are sorted in various ways such as by seasonal average, by annual average, by elevation above the terrain, or by land topography. State wind resource maps recently produced by NREL show much greater detail than the national map shown in Fig. 11-8.

The fact that the atmosphere has viscosity and that the lowest layers of the wind are being slowed by contact with the earth's surface means that the wind speed will change according to the altitude above the earth's surface. Although this wind shear depends on many local variables, the commonly accepted empirical expression for this change below about 300 m is a power law:

$$v_2/v_1 = (h_2/h_1)^{\alpha} \tag{11-11}$$

where α = about a one-seventh average, v_n = wind speeds at vertical locations 1 and 2, and h_n = elevations above a reference plane for 1 and 2.

Although the wind by its nature is stochastic, for a given geographical location, it is possible to assign a characteristic wind speed probability density curve. Such curves, together with power curves for a particular model of wind machine, can be used for good estimations of the annual energy that can be captured and therefore a good estimation of annual revenue from a proposed wind power plant. Because the power that can be extracted from the wind varies as the cube of the instantaneous wind speed, a region having, for example, a 2% higher annual mean wind compared to another region will be expected to yield 6% more energy than the second region, all other things being equal. This difference can be financially important for a wind farm operator.

Many probability expressions have been examined, but present usage has evolved to just a few. For example, a Gaussian curve will usually give a very good fit to a 10-min wind speed data set. However, the Weibull expression is more widely accepted for annual records. Two parameters, c and k, give the Weibull curve its flexibility to more nearly fit experimental data. The c parameter adjusts the numerical scale on the abscissa or independent variable axis. The shape factor, k, adjusts the relative width of the peak of the curve:

$$f(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{(k-1)} e^{-\left(\frac{v}{c}\right)^k} \tag{11-12}$$

Examples of the Weibull function are given in Fig. 11-9 for $k = 2$ and $k = 3$.

Good energy sites for wind machine installations very often can be fit with a shape factor of $k = 2$. This special case of the Weibull expression is the well-known Rayleigh function. A useful theoretical model can be made by postulating a perfect wind machine that can operate in all winds at the Betz limit of power coefficient, namely $0.593 = {}^{16}/_{27}$. Assume this machine is placed in a region where the wind statistics exactly fit a Rayleigh probability density function. Calculations based on this model can be used to set upper limits to the amount of annual energy that can be captured by any real machine in a similar location. In particular, if inertia effects are ignored, and the product of this ideal Betz power curve and the Rayleigh function is integrated over all winds, the result will be the mean annual power of a "Rayleigh-Betz" wind machine. It can be reduced to the following expression[5]:

$$P_{av} = \rho\,({}^2/_3\,D)^2\,V_{av}{}^3 \tag{11-13}$$

where ρ is air density, D is the diameter of the swept area, and V_{av} is the mean annual wind. The implied annual energy that this ideal model will collect will therefore be the product of the hours in a year with this average power. For example, if a wind machine with a 15-m rotor diameter is operated at sea level for a year in an ideal Rayleigh wind regime of mean value 5 m/s, then the energy captured could not exceed

$$\begin{aligned} W &= 8{,}760 \text{ h/year} \times 1.2 \text{ kg/m}^3 \times ({}^2/_3 \times 15 \text{ m})^2 \times (5 \text{ m/s})^3 \\ &= 131.4 \times 10^6 \text{ watthours} = 131.4 \text{ megawatthours} \end{aligned}$$

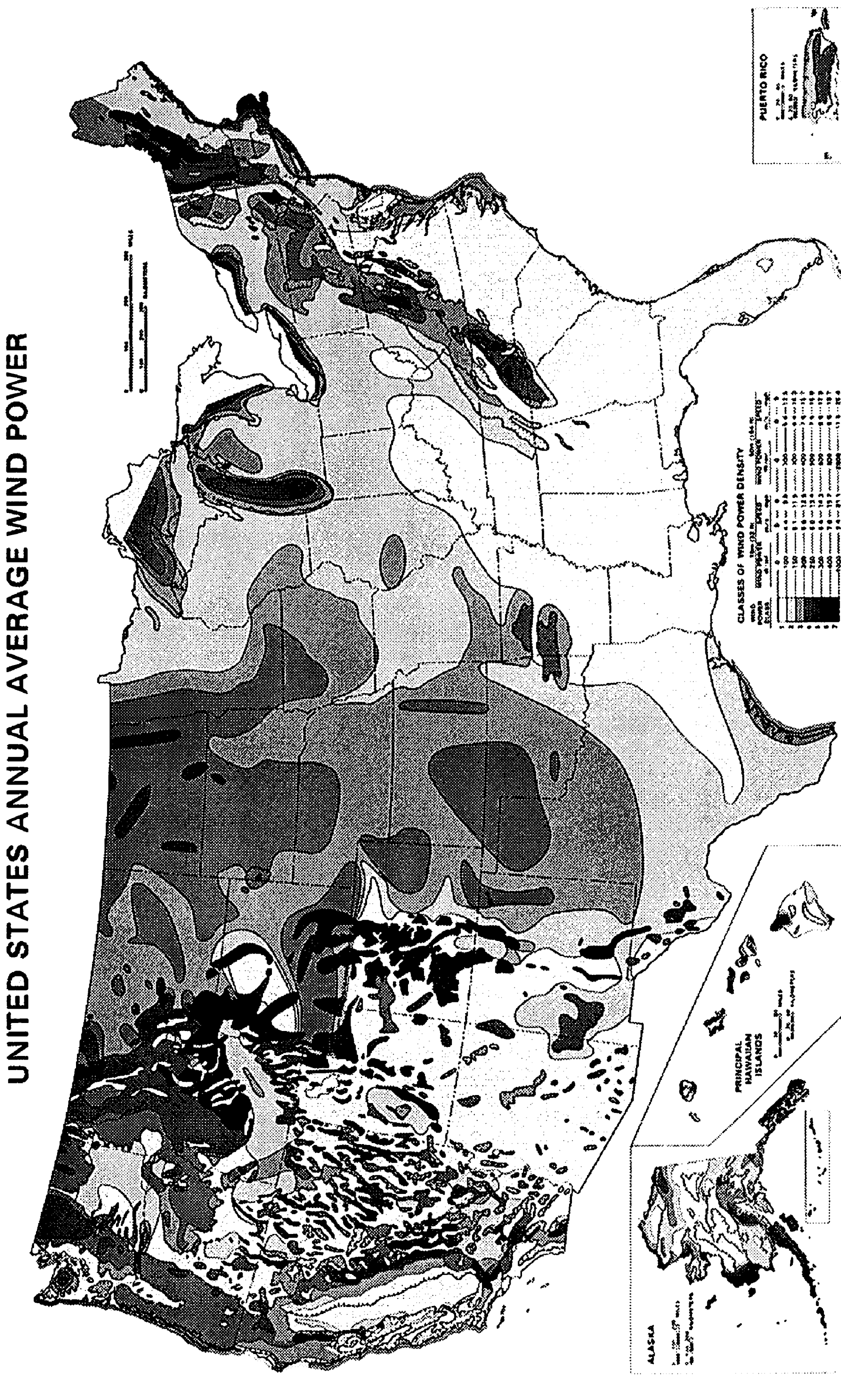


FIGURE 11-8 Wind energy resource map of the United States.

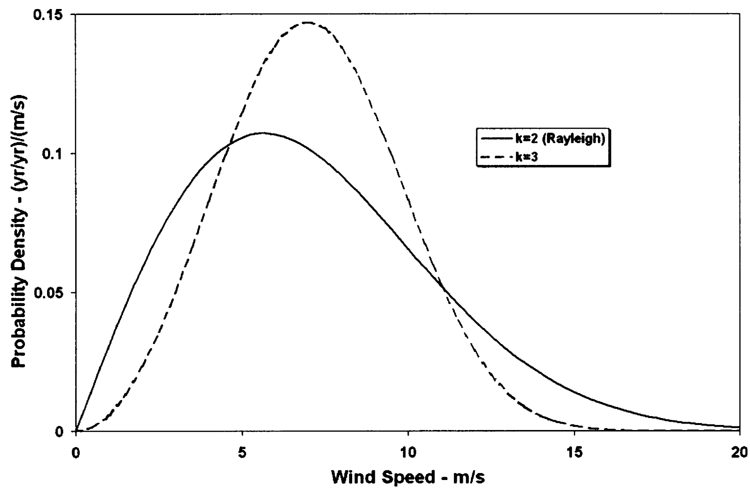


FIGURE 11-9 Weibull probability densities for scale factor $c = 8$ m/s.

Note that average air density has an important influence on calculated wind turbine power, as shown in this example: The sea level air density of 1.2 kg/m^3 falls off to 1.0 kg/m^3 at about 1,500 m above sea level. This of course means a power reduction back to 83% of sea level power for the same wind speed. If seasonal changes in temperature and barometric pressure are known, corrections to annual calculated power can be made.

Being able to estimate annual average wind speeds and annual output is useful for planning purposes, but being able to forecast the wind accurately for one or two days could provide a significant economic benefit, allowing utility engineers to plan ahead for wind generation to replace fossil-fueled generation. Various methods have been developed, and are continuing to be developed, to accurately forecast wind power one to two days ahead. Several commercial companies offer this wind forecasting service.

11.5.7 Wind Turbine Electric Systems

Generators. During the rapid development of wind power plants in the late 1980s and early 1990s, nearly all grid-connected wind turbine generators were of the induction type. Induction generators are much less expensive than the synchronous type, but their primary flaw is their need to draw lagging current from the grid to supply their excitation. This apparent disadvantage is sometimes useful for a utility in that if the utility to which they are paralleled goes down, induction generators loose their reactive excitation from the utility grid and automatically cease producing power. Presumably, aerodynamic controls on the wind turbine rotor will limit its speed if it loses its load in this manner.

It has become commonplace for electric utilities to require induction generator owners to provide power factor correction capacitors that will supply nearly all of the required lagging current. The flaw of this solution is that if the interconnected utility were to have a forced outage, there is the possibility for self-excitation of the remaining online wind generators. Thus, if this isolated subsystem happens to have generation approximately equal to its load, then the system could sustain itself long enough to cause considerable damage. At present, this scenario is essentially impossible because of the wind turbine's control system, which halts the turbine if either frequency or voltage stray beyond given narrow limits.

At present, a growing electric generator technology receiving consideration for wind systems is the permanent magnet (PM) type. Certain small wind turbine systems in the range of less than 10 kw have employed such generators for many years. Progress in the development of rare earth permanent magnets and other types of permanent magnets has made the use of PM generators more feasible for larger-sized wind systems. The advantages of PMs over conventional synchronous generators are simplicity of construction and no energy loss in the field winding. These advantages are achieved at the price of no adjustment of field strength and the high cost of permanent magnets. The loss of flexibility in the PM generator can be largely compensated for with semiconductor power processing.

Power Electronics. The continuing progress in the development of larger and less expensive semiconductor devices has opened the way to variable-speed, constant-frequency wind energy systems. The rapid growth of the use of adjustable speed drives for motors in manufacturing production lines illustrates the present high reliability of power electronics. The two most frequent methods of using this technology for generation are (1) conversion of the entire variable-voltage, variable-frequency output of a wind machine (sometimes called "wild ac") to direct current, then reconverting it to utility quality ac power; and (2) alternatively, when a wound rotor, variable-speed induction machine is fed with an appropriate variable frequency current from a power semiconductor system, the stator will supply constant voltage, constant frequency power to the interconnected utility bus. This arrangement is usually called the doubly fed induction generator and offers the advantage of a smaller power rating for the power electronics equipment.

11.5.8 Controls and Control Algorithms

The prime movers for conventional utility grid generators, having been developed along with their companion generators, are capable of very precise speed and power control. These levels are totally under the control of the utility dispatcher.

Wind turbines, however, must exploit whatever wind presents itself at a particular instant. Anemergonic (wind energy) engineers have agreed on a taxonomy to describe the three successive operating regions that a wind turbine progresses through as the ambient wind increases from calm to maximum. Region one is that region preceding startup and below cut-in. Cut-in is the point at which there is just sufficient wind to produce measurable energy. For most machines this is in the range of 4.5 to 5 m/s.

Region two identifies operation between cut-in and rated power operation; it is where the electrical power output uniformly increases as wind speed rises from cut-in speed. For constant-speed wind turbines in this region, rotor speeds are selected such that the maximum power coefficient will occur when the most productive wind exists. The most productive wind is that speed, which if a wind turbine were allowed to operate only in a narrow band centered on this speed, would capture more energy than being centered on any other narrow band of wind speed. This wind speed exists because of the opposing effects of exploitable energy rising with the cube of wind speed and the decline in frequency of occurrence of such winds for increasing speeds. For the special case of a wind regime governed by the Rayleigh probability density, this wind speed occurs at 159% of the annual average wind and at exactly twice the most probable wind speed. The range of possible values for the wind speed that defines the upper edge of region two is quite broad and is machine dependent. It is usually the least wind that will yield the maximum continuous power output for that turbine.

Variable-speed wind turbines in region two attempt constantly to adjust their speed to hold their TSR at the maximum power coefficient point. This flexibility is purported to make it possible to collect theoretically up to 20% more annual energy than the same machine constrained to operate at constant speed. The obvious consideration is to balance this gain with the expense and power losses in any auxiliary equipment such as power electronics modules that are needed to provide the variable-speed feature.

As steady-state operation in wind speeds beyond rated speed would overheat the generator, means must be taken to limit wind power input. This identifies the upper edge of region two, which is the lower edge of operating region three. It is the region where the control algorithm must change from attempting to maximize energy capture to minimizing equipment damage. Although in principle, generator torque can be used to continuously limit wind machine speed to its specified maximum value, the well-designed wind machine always has an additional means of aerodynamic control to prevent a utility failure from allowing the wind turbine to accelerate out of control. Although it is at the cost of appreciable mechanical complication of the rotor hub, the most common and effective aerodynamic control is that of continuous control of pitch angle of the wind machine rotor blades.

Even though there is only one best angle for a particular wind machine to operate at as can be seen from the surface in Fig. 11-7, the figure also shows that increasing blade pitch can rapidly reduce power coefficient and therefore output power from the rotor. In fact, at 90° pitch the rotor should, in principle, come to rest. In real applications, especially in larger machines, there is the practical problem of pitching speed and whether blades can be pitched quickly enough to achieve the appropriate power or mechanical load reduction.

Design and testing of control algorithms that minimize fatigue damage of a wind turbine drive train and blades without lessening energy capture continues to be an important area of research.

11.5.9 Computer Simulation

Having reviewed energy collection, electrical generation, and dynamic control of wind turbines, it is not surprising that there are in general three families of computer simulation programs that provide insight into some of these areas.

There are the performance programs that simulate energy capture for hybrid electric systems. A *hybrid system* in an energy production system that combines two or more energy sources such as wind, solar, and diesel. Using either typical time history wind samples or a wind probability density curve together with a typical electrical load history, these programs will forecast the energy from the wind turbines together with the run time of any associated engine generator set. Time periods of interest are not less than months.

The next class of simulation is focused on aerodynamics, mechanical dynamics, and structures. Solutions yielded include such things as static and varying loads on structures, predicted power curves, normal vibration modes of machine components, and control system responses. Time scales treated are from seconds to minutes.

The last category includes the electrical simulations of the transients in generators and power electronics. Simulation results can warn of possible self-excitation of induction generators, overvoltages due to harmonic resonances from power electronic modules, or excess starting transient currents. Time scales used could be from microseconds to a few seconds.

Figure 11.10 illustrates three types of investigations that require different time scales. Electrical short circuits and resulting transients occur in a few seconds, while voltage regulation occurs over minutes and load following up to a day.

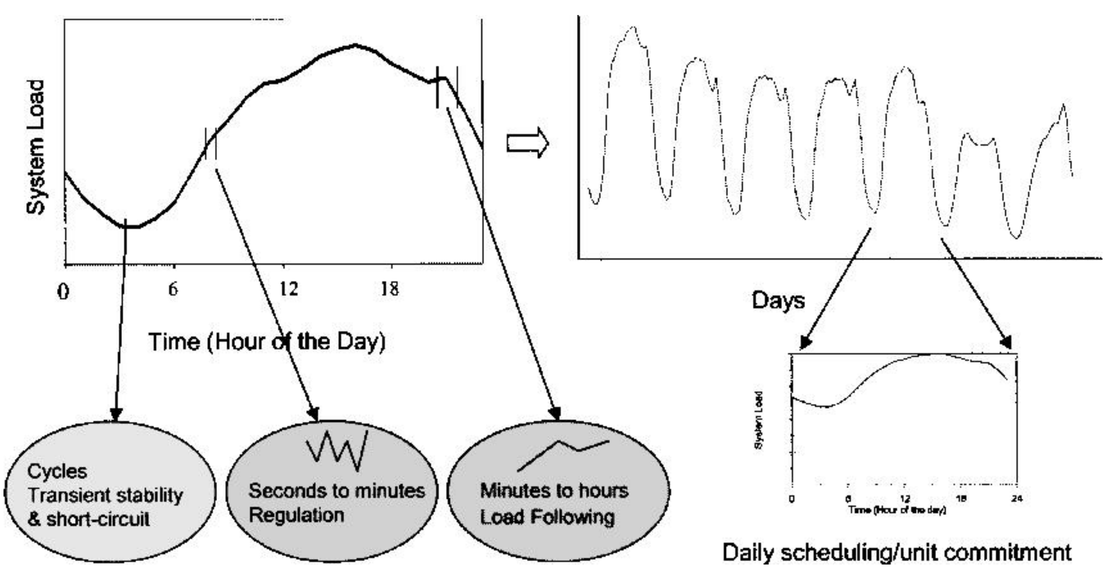


FIGURE 11-10 Three Simulation Regimes.

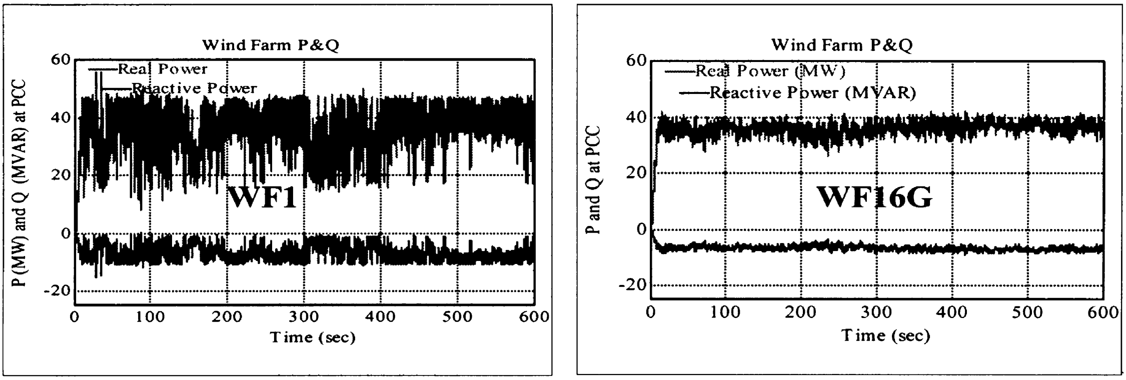


FIGURE 11-11 Real and reactive power output of a wind power plant with different aggregations.

Wind power plant designers have rightfully designed turbine controllers and safety features to disconnect and stop a turbine if it detects a fault such as a short circuit in the connection to the utility. Unfortunately, in practice, when this occurs at a wind farm of up to 50 turbines, they usually will all shut down even if the fault clears instantaneously. Each turbine must then be individually restarted. This interesting problem is called *low voltage ride through* (LVRT), and is used to describe the wind turbine's ability to stay connected during such a disturbance. The solution to this situation must be a balance between safety and inconvenience.

Figure 11-11 is an example of simulation that shows the effect of spatial diversity in a wind plant and how it naturally smoothes the power delivered.

Figure 11-11 shows a simulation of the real and reactive power of a wind farm of 200 turbines with a total rated output of 45 MW. This figure compares two methods of aggregation. The left graph shows the output of the simulation when we assume that the entire wind farm is represented by identical wind turbines, all turbines are located at the same spot and each one of them is driven by the same wind speed. Thus, the collective output of individual turbine is reflected as the output of the entire wind farm at the point of interconnection. In the right graph, the wind farm simulation groups the wind turbines into 16 different groups to simulate the diversity of a wind farm. Each group has a different number of wind turbines and is driven by a time series wind speed. Each series is time shifted with respect to other wind speed files driving different groups of wind turbines. The time shift approximates the time delay of the wind reaching different groups of wind turbines caused by differences in geographical location among the groups. Although this method is not perfect, it is closer to the reality of a wind farm. It clearly shows that the power fluctuations at the point of interconnection are significantly reduced when we consider the aggregation for 16 groups of turbines because of cancellation of the power fluctuations among the groups.

11.5.10 Issues Related to Wind Turbine Use

By Charles Butterfield, Michael Millian, Eduard Muljadi

Classes of Usage for Modern Wind Turbines. Commercially available wind turbines for electric power range in size from less than 100 W to a several megawatts. Although today wind turbines are employed in many ways, a few important classes of applications can be identified. The most publicized and largest in terms of aggregate power is, of course, the wind power plant application. Modern wind power plants can be as large as 300 MW and are often located within a short distance of each other. These wind power plants feed the same utility line. Although there are some significant policy drivers that have influenced the development of wind power plants, today wind is often economically competitive with conventional generation. It is therefore likely that significant new wind generating capacity will be developed in the next several years.

Another class called "wind-hybrid" denotes one or more midsized turbines operating in conjunction with other power sources such as a diesel generator, a photovoltaic system, and other sources of power that are all managed by a common control system. Wind-hybrid systems provide a means to reduce the costs of electric power in outlying regions where conventional fuel is expensive and difficult to transport. Remote villages in developing countries and polar regions are the usual applications of this class.

Small turbines comprise the third class of wind turbine. They usually are privately owned for use on ranches, farms, vacation cabins, and remote communication repeater stations.

The remainder of this section is a collection of several unrelated topics that nevertheless affect the employment and acceptance of wind energy.

Certification. As in many other areas of technology, prospective buyers need to have confidence in the claims of a manufacturer. This leads to the concept of certification such as those that exists in the field of electrical appliances. The United States is participating in the development of International Electrotechnical Commission (IEC) standards for wind turbines through the American Wind Energy Association (AWEA) and the American National Standards Institute (ANSI) memberships in IEC. The IEC 61400-XX series covers design requirements, testing requirements, and certification guidelines. At present there are two dominant international certification bodies active in the European wind turbine industry: Germanischer Lloyd[7] and Det Norske Veritas (DNV). Each certification body has developed its own rules and requirements for type certification, but they generally have the same goal: to review the design documentation and ensure that the turbine has been designed with engineering discipline in accordance with recognized industry standards, and that it will function reliably throughout its specified life. Most turbines installed in Europe must be certified by one of these bodies as a national requirement. Although the United States does not require such certifications, yet to alleviate financial risks, most commercial-scale turbines are certified.

These international certifications do not satisfy utility interconnection requirements in the United States. Interconnection requirements are usually set by individual utilities for large-scale turbines. However, the Federal Energy Regulatory Commission (FERC) is attempting to set general interconnection rules such as FERC Order 661, Interconnection of Wind Turbines. For small turbines, IEEE 1547 defines interconnection requirements that can be certified by Underwriters Laboratories (UL) and other certification bodies and are usually accepted by local electrical inspectors.

Over the past 5 years, the European community has encouraged harmonization of the different certification rules to facilitate trade within Europe. This has resulted in active standards development programs on the international level, primarily IEC.

The Wind Turbine Research Program conducted for DOE by NREL employs a comprehensive engineering development process that includes regular design, testing, and documentation reviews throughout the process. This process follows accepted international procedures, including the International Standards Organization—ISO 9001.

Grid Integration and Operational Impacts. Wind power plants have an impact on power system operations. Generally, these impacts can be divided into timescales that correspond to utility operational practice (see. Fig. 10). The regulation timescale is from a few seconds to several minutes. In the regulation, timescale generation responds to small, fast fluctuations in the load, primarily via automatic generation control (AGC) computers. These fast fluctuations are generally uncorrelated.

The next load following timescale spans several minutes to several hours. Load fluctuations in this timescale are managed by generating units that are economically dispatched. Changes in load over this timescale are more highly correlated than in the regulation timescale and fluctuate over a wider range. *Unit commitment* is the process of determining which of the slow-start (thermal) units will be needed for the day-ahead schedule and ensuring those units are available when needed. Wind power plants generally increase the need for regulation and load following service and can complicate the unit commitment process. Over the planning horizon, considerations of system adequacy can be evaluated using reliability models. System adequacy is often measured by loss of load probability or a related metric. The contribution of wind power plants to system adequacy is often measured as the effective load carrying capability (ELCC) of the wind power plant. Because of the intermittent nature of the wind, ELCC as a percent of rated capacity is generally low and often falls in the range of 10% to 40% of rated capacity. Several power pools, regional transmission organizations, and other entities have developed simple approaches to calculating wind capacity value. These are detailed in Ref. 8.

TABLE 11-4 Some Samples of Integration Costs to Utilities of Wind Power Plants

Study	Relative Wind Penetration (%)	Regulation ($/MWh)	Load Following ($/MWh)	Unit Commitment ($/MWh)	Total ($/MWh)
UWIG/Xcel	3.5	0	0.41	1.44	1.85
PacifiCorp	20	N/A	1.64	3.00	4.64
BPA/Hirst	7	0.19	0.28	1.00–1.80	1.47–2.27
PJM/Hirst	0.06–0.12	0.05–0.30	0.70–2.80	N/A	0.75–3.10
We Energies I	4	1.12	0.09	0.69	1.90
We Energies II	29	1.02	0.15	1.75	2.92
Great River Energy I	4.3				3.19
Great River Energy II	16.6				4.53
CA RPS Phase I	5	0.17	0	N/A	0.17
MN DOC/Xcel	15	0.23	0	4.37	4.60

Several detailed studies have been performed to analyze the impact of wind power on system operations. The Utility Wind Interest Group (UWIG) is an organization that serves as a clearinghouse for challenges and solutions related to wind energy, and UWIG was the sponsor of one of the early studies. The general approach of the recent wind integration studies is to perform a full system analysis using standard software tools, recognizing that wind contributes an additional source of variability into an already variable system. Because the system needs to be balanced in aggregate, it is not necessary to balance each movement in wind one-for-one by a counter move of a conventional generator, and the standard simulation tools perform a system optimization that observes the requirements for balancing loads and resources. One summary of studies performed in the United States appears in Ref. 7. To analyze the impact that wind has in the regulation timescale, a method developed at the Oak Ridge National Laboratory is often used[9]. This approach allocates the regulation burden to variable loads and resources, and has been applied to wind in several analyses. Table 11-4 has been updated to reflect PacifiCorp's 2004 Integrated Resource Plan[10]. The table shows the range of integration costs from recent studies of wind's impact on power system operation and economics.

11.5.11 System Operation with Wind Power

By Yih-Huel Wan

When a wind power plant is connected to the grid, power and voltage fluctuations are concerns for system operators. Modern wind turbines can provide very good voltage regulations, but the power from the wind power plant is not controlled by its operators. However, actual wind power plant output data show that although wind power fluctuations are stochastic in nature, they are not completely random. Because of wind speed persistency, the magnitudes and rates of power level changes from wind power plants are seldom extreme. Wind speed does not change instantaneously, nor does power from wind power plants.

Short-time step changes of wind power are small. For example, the average magnitude (absolute value) of second-to-second step changes is less than 0.2% of the nameplate capacity for small wind power plants (with tens of turbines) and 0.1% for large wind power plants (with hundreds of wind turbines). The average magnitude of minute-to-minute step changes ranges from about 1.2% of the nameplate capacity for small wind power plants to about 0.3% of the capacity for large wind power plants. Power from large wind power plants is less variable than power from small wind power plants because of spatial variation of wind speeds. Physical separations and differences of local terrains cause wind speeds to vary. Even adjacent wind turbines within the same wind power plant do not experience the same wind conditions during short time frames and their instantaneous outputs are not likely to be synchronized (e.g., outputs from some turbines are increasing while others are decreasing). Such spatial variation in wind speeds makes the combined outputs from more turbines less volatile, especially in short time frames.

Wind power can experience bigger changes during longer periods, but average magnitudes of wind power changes for longer time steps are still relatively small. Average magnitudes of 10-min step changes range from 3% of the nameplate capacity for small wind power plants to 2% of the nameplate capacity for large wind power plants, while average magnitude of hourly step changes ranges from 7% to about 5% of the nameplate capacity.

In addition to small average magnitude of wind power step changes, the standard deviation values of wind power step changes are also small. This indicates that most of the wind power step changes are of small values and large step changes are relatively rare. As an example, Fig. 11-12 shows the distribution of 10-min wind power step changes for 1 year from a large wind power plant with a nameplate generating capacity of 100 MW. Over 98% of all 10-min step changes are within $\pm 3\sigma$ (or $\pm$11 MW or $\pm$11% of the power plant's nameplate capacity in this case). Large swings of wind power do occur, but those infrequent large changes are almost always related to well-defined weather events such as storms and weather fronts. Most of these weather events can be accurately predicted in advance to allow system operators time to take remedial actions.

Step change statistics define the outer boundary of the wind power fluctuations. The rate of change of wind power levels in a given time interval (e.g., 10 min or 1 h) is another indicator of wind power plant behavior. Power levels from a wind power plant fluctuate continuously because the wind speed changes continuously as it moves through the wind power plant. For example, in a 1-s wind power data series, the average duration for wind power increases or decreases is only 2.2 s (with a standard deviation of 4.4 s). With 1-min data series, the average duration of the increases and decreases is 2.4 min (with a standard deviation 8.4 min). Consequently, wind power plant ramping rates are also very modest. Table 11-5 lists the wind power 10-min ramping statistics.

As shown in the table, average ramping rates are less than 0.5% of the power plant's nameplate capacity per minute. The maximum ramping rates can be large, especially for the negative ramping (rapid wind power decreases). However, such big changes occur infrequently. The distribution of ramping rates of three wind power plants for 1 year is shown in Table 11-6. It shows that wind power plant ramping rates seldom exceed 5% of the nameplate capacity per minute.

Further examination of the data indicates that not all of those very large ramping rates are the result of wind speed changes. Again, most of these large ramping rates are related to well-defined weather events that can be forecasted.

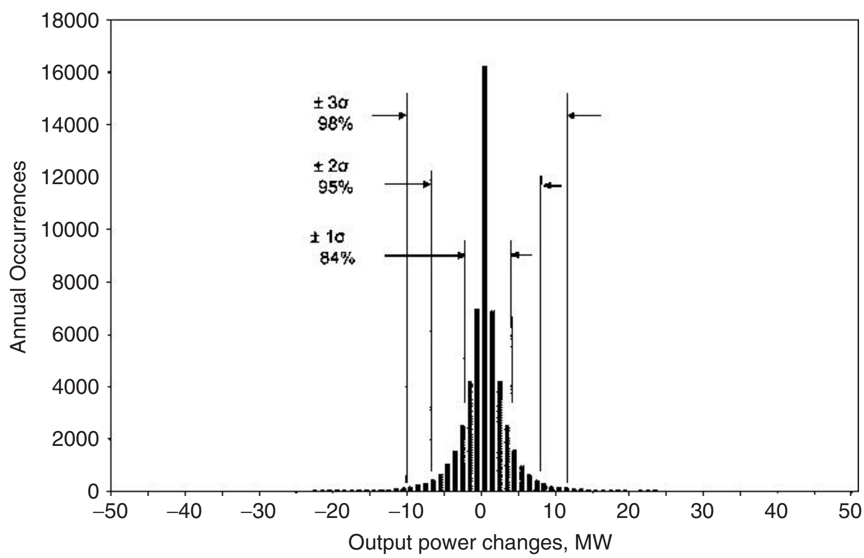


FIGURE 11-12 Wind plant power output changer during 10-min interval.

TABLE 11.5 Wind Power 10-min Ramping Characteristics

	Average (kW/min)	Average (%/min)	Standard deviation (kW/min)	Standard deviation (%/min)	Max (+) (MW/min)	Max(−) (MW/min)
10 MW Plant	47	0.45%	78	0.74%	1.2	–1.4
100 MW Plant	269	0.26%	465	0.45%	7.0	–12.2
220 MW Plant	438	0.18%	811	0.34%	11.3	–29.0

Actual operations experience suggests that most control areas will have few problems integrating wind power at relatively low penetration levels. Utilities in several Midwestern and western states have had wind power supplying from 6% to 14% of their online loads during high wind periods. Recent studies also concluded that impacts of wind power on system operations are minor.

11.5.12 Wind Turbine Acoustic Noise

The importance of sounds emanating from a wind turbine obviously depends on its surroundings. In remote locations of the deserts of California or Texas or in offshore installations worldwide, the relatively low-level sound emitted by a group of wind machines is unimportant. In northwestern Europe, where a single machine may stand near a cluster of houses, even low-level sounds, especially at night, can be annoying. Wind turbine noise can be divided into two classes: mechanical and aerodynamic. Mechanical noise includes transmission gear tones and generator whine. These are common machinery problems and can be ameliorated by the usual methods of gear selection and soundproof insulation.

Aerodynamic noise, in turn, can be divided into impulsive and broadband noise. Impulsive noise essentially always comes from a downwind rotor blade passing behind its tower; it is caused by the sudden change in wind flow across the rotor blade. This can be modified by spiral strakes and other shapes attached to the tubular tower or changing the blade-to-tower distance. For lattice towers, the effect is somewhat less but is still present. VAWTs also have a shadow due to the central supporting column, but it is somewhat less pronounced than for HAWTs.

Broadband noise is the swishing noise of the air interacting with the turbine blades and is generated in varying degrees by all turbines. There are obviously many sources for this sound, but three important ones are blade tip vortex, blunt trailing edge airfoil, and turbulence in the wind itself. At the present writing, designers agree that broadband noise decreases with decreasing linear tip speed.

For more information on wind turbine acoustics, refer to Ref. 11.

11.5.13. Wildlife Considerations

By karin Sinclair

As wind turbines are installed across the country, concerns over possible wildlife impacts have expanded. Throughout the 1990s, the principal concern was for avian impacts. In 1992, primarily because of possible negative impacts caused by wind power development on golden eagles (Aguila chrysaetos) in the Altamont Pass Wind Resource Area in California, DOE and NREL, working

TABLE 11-6 Distribution of Ramping Rates

	$< \pm1\sigma$ (78 kW/min)	$< \pm2\sigma$ (156 kW/min)	$< \pm3\sigma$ (234 kW/min)	$< \pm4\sigma$ (312 kW/min)	$< \pm5\sigma$ (390 kW/min)	> 5%/min (525 kW/min)
10 MW Plant	0.82476	0.95146	0.98305	0.99299	0.99673	0.00122 (64)
	$< \pm1\sigma$ (0.5 MW/min)	$< \pm2\sigma$ (1.0 MW/min)	$< \pm3\sigma$ (1.5 MW/min)	$< \pm4\sigma$ (2.0 MW/min)	$< \pm5\sigma$ (2.5 MW/min)	> 5%/min (5.2 MW/min)
100 MW Plant	0.84047	0.95742	0.98450	0.99304	0.99619	0.00040 (21)
	$< \pm1\sigma$ (0.8 MW/min)	$< \pm2\sigma$ (1.6 MW/min)	$< \pm3\sigma$ (2.4 MW/min)	$< \pm4\sigma$ (3.2 MW/min)	$< \pm5\sigma$ (4.0 MW/min)	> 5%/min (12.1 MW/min)
220 MW Plant	0.86471	0.96779	0.98856	0.99438	0.99649	0.00044 (23)

collaboratively with stakeholders including utilities, environmental groups, consumer advocates, utility regulators, government officials, the wind industry, and the National Wind Coordinating Committee's (NWCC) Avian Subcommittee, formed an active avian-wind power research program.

Avian concerns fall within two main areas: the effect of avian mortality on bird populations, and possible litigation over the killing of even one bird if it is protected by the U.S. Migratory Bird Treaty Act or the U.S. Endangered Species Act or both. After a decade of research, focused predominantly on developing solutions to reduce or avoid avian mortality caused by wind energy development throughout the United States, the DOE/NREL research task came to an end. A list of the reports generated from these research projects can be found at http://www.nrel.gov/wind/avian_reports.html.

In 1999, the Avian Subcommittee of the NWCC published guidelines for conducting avian research (Studying Wind Energy/Bird Interactions: A Guidance Document). These guidelines contain metrics and methods that should be applied to all current avian research projects. It was anticipated that the use of a standardized set of metrics and methods would help facilitate comparability among research sites.

Since DOE/NREL ended its research on issues related to avian impacts, concerns for other wildlife issues have emerged. For example, impacts on bats have become a concern as a result of high bat fatalities found at two newer wind power plants in the northeast (the Mountaineer Wind Energy Center in Tucker County, West Virginia, and the Meyersdale Wind Energy Center in Somerset County, Pennsylvania). In 2003, bat carcasses were found while conducting traditional post-construction avian fatality surveys. Although bat carcasses have been found at other wind power plants across the country, the number of fatalities found at these two sites far exceeds anything found to date. Bat-specific fatality searches have now begun at other U.S. wind power plants.

Two other issues, potential impacts on nocturnal species and grassland/shrub steppe species, have been identified by the NWCC's Wildlife Workgroup (previously known as the Avian Subcommittee). To begin addressing these issues, the Wildlife Workgroup will develop a companion document to the Guidance document, focusing on nocturnal issues such as nocturnal behavior of birds and bats. A literature review of wind power impacts on grassland/shrub steppe habitats will also be produced.

11.5.14 Summary

Over the past 30 years, wind energy has passed through its infancy and childhood and is now achieving a measure of maturity as is evidenced by the turning of researcher efforts from physics and performance of wind turbines to how energy derived from wind can better interface with today's technology and be accepted for what it can contribute. For example, the committee on wind energy within the International Energy Association now supports special groups studying *integration of wind and hydropower, wind energy in cold climates,* and *dynamic models of wind power plants for power system studies*.

By January 2006, the world had an installed capacity of 56 GW, and growth is expected to continue at 28% per year for a few years. In some countries, wind energy is supplying up to 20% of that nation's electrical power.

According to the IEA Annual Report, traditional technologies average electrical generation costs are in the range $25 to $45 per megawatthour. Levelized wind generation ranges from $35 to $95 per megawatthour.

BIBLIOGRAPHY

References

1. American Wind Energy Association, 122 C Street, NW, 4th Floor, Washington, DC 20001, (202) 383-2504, http://www.awea.org/.
2. *The Implementing Agreement for Co-operation in the Research and Development of Wind Turbine Systems within The International Energy Agency*, PWT Communications, 5191 Ellsworth Place, Boulder, CO 80303, http://www.ieawind.org/.
3. Lanchester, F.W., "Contributions to the Theory of Propulsion and the Screw Propeller," *Transactions of the Institution of Naval Architects,* vol. LVII, March 25, 1915, pp. 98–116.

4. Betz, A., "Das Maximum der theoretisch möglichen Ausnützung des Windes durch Windmotoren," *Zeitschrift für das gesamte Turbinenwesen,* Heft 26, Sept. 26, 1920.
5. Manwell, J. F., McGowan, J. G., Rogers, A. L., "*Wind Energy Explained Theory, Design and Application,*" John Wiley & Sons, New York, 2002.
6. Elliot D. L., Holladay, C. G., Barchet, W. R., Foote, H. P., and Sandusky, W. F., *Wind Energy Resource Atlas of the United States,* DOE/CH 10093-4, Golden, CO, Solar Energy Research Institute, March 1987. Available at http://www. rredc.nrel.gov/wind/pubs/atlas/.
7. Parsons, B., Milligan, M., Zavadil, B., Brooks, D., Kirby, B., Dragoon, K., and Caldwell, J., "Grid Impacts of Wind Power: A Summary of Recent Studies in the United States," in *Proceedings of the European Wind Energy Conference and Exhibition*, Madrid, Spain, European Wind Energy Association, NREL/CP-500-34318, June 16–19, 2003. Available at http://pix.nrel.gov:8020/BASIS/nich/www/nrel/SDF.
8. Milligan, M., and Porter, K., "Determining the Capacity Value Of Wind: A Survey Of Methods And Implementation," in *Proceedings of WindPower 2005*, May 15–18, 2005. Denver, CO, American Wind Energy Association, NREL/CP-500-38062.
9. Kirby, B., and Hirst, E., "*Customer-Specific Metrics for the Regulation and Load Following Ancillary Services*." Oak Ridge National Laboratory, Oak Ridge, TN, 2000.
10. Smith, J., DeMeo, E., Parsons, B., and Milligan, M., "Wind Power Impacts on Electric-Power-System Operating Costs: Summary and Perspective on Work to Date," in *Proceedings of WindPower 2004*, Chicago, IL, American Wind Energy Association, NREL/CP-500-35946, March 29–31, 2004. Available at http://www.nrel.gov/docs/fy04osti/35946.pdf.
11. Wagner, S., Bareiss R., and Guidati, G., *Wind Turbine Noise*, ISBN 0-387-60592-4, Springer-Verlag, Heidelberg, Berlin, 1996.

Additional Information Sources

Bergey, K. H., "The Lanchester-Betz Limit," *J. Energy*, vol. 3, Nov.–Dec. 1979, pp. 382–384.

Betz, J. A., *Introduction to the Theory of Flow Machines*, Pergamum, New York, 1966.

Eggleston, D. M., and Stoddard, F. S., *Wind Turbine Engineering Design*, Van Nostrand Reinhold Company, New York, 1987.

EnerNex Corporation and Wind Logics, Inc., "Wind Integration Study—Final Report." Xcel Energy and the Minnesota Department of Commerce, 2004.

GE Energy Consulting, "The Effect of Integrating Wind Power on Transmission System Planning, Reliability, and Operations Report on Phase 2: System Performance Evaluation," The New York State Energy Research and Development Authority, 2005.

Germanischer Lloyd web page: http: //www.germanlloyd.de/Activities/Wind/windener.html

Gipe, P., *Wind Power for Home and Business*, Post Mills, VT, Chelsea Green Publishing Company, 1993.

Johnson, G. L., *Wind Energy Systems*, Prentice Hall, 1995.

Spera, D. A., ed., *Wind Turbine Technology*, New York, ASME Press, 1994.

Troen, I., and Petersen, E. L., *European Wind Atlas*, Published for the Commission of the European Communities Directorate-General for Science, Belgium, by Risø National Laboratory, Roskilde, Denmark, 1989.

Wan, Y., "Fluctuation and Ramping Characteristics of Large Wind Power Plants," proceedings of American Wind Power Association WindPower 2005 Conference, May 15–18, 2005.

Wan, Y., Milligan, M. and Parsons, B., "Output Power Correlation Between Adjacent Wind Power Plants," Journal of Solar Energy Engineering, Transactions of the ASME, vol. 125(4), November 2003, pp. 551–555, NREL Report No. JA-500-33519. Available at http://www.pix.nrel.gov:8020/BASIS/nich/www/nrel/SDF.

Wan, Y., and Bucaneg, D., "*Short-Term Fluctuations of Large Wind Power Plants,*" 2002, NREL CP 500-30747. Available at http://www.pix.nrel.gov:8020/BASIS/nich/www/nrel/SDF.

Periodical Publications and Reports

Windpower Monthly, P.O. Box 4258, Grand Junction, CO 81502-4258.

Wind Energy Weekly, AWEA, 122 C Street , NW, 4th Floor, Washington, DC 20001.

Annual Reports of the Implementing Agreement for Co-operation in the Research and Development of Wind Turbine Systems of the International Energy Agency. Published by the National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO, 80401-3393.

WindStats Newsletter, P.O. Box 100, 8420 Knebel, Denmark, http://www.windstats.com

Federal Wind Energy Program

U.S. Department of Energy Wind Energy Program, http://www.eere.energy.gov/windandhydro/.

National Renewable Energy Laboratory, National Wind Technology Center, http://www.nrel.gov/ wind/.

Sandia National Laboratories, P.O. Box 5800, Mail Stop 0708, Albuquerque, NM 87185-0708, http://www.sandia.gov/Renewable_Energy/.

Wind Energy Organizations

The American Wind Energy Association (AWEA), 1101-14th St, NW, 12th Floor, Washington, DC 20005, (202) 383-2500, http://www.awea.org

Utility Wind Interest Group, http://www.uwig.org

National Wind Coordinating Committee, http://www.nationalwind.org

Conferences

Two major wind energy conferences are held every year: Windpower and the Wind Energy Symposium. Windpower is sponsored by AWEA; the Wind Energy Symposium is held in conjunction with the AIAA Aerospace Sciences Meeting & Exhibit. Both conferences publish proceedings.

11.6 GEOTHERMAL POWER

BY RAYMOND FORTUNA

The outer crust of the earth contains a very large reservoir of energy present as sensible heat. This resource is between one and two orders of magnitude larger than the recoverable energy from uranium and thorium in the same volume of rocks. If 1% of the thermal energy contained within the uppermost 10 km of the earth could be tapped, the amount of heat would be 500 times that contained in all oil and gas resources of the world. The only natural fuel system that represents a larger energy resource on the earth is the fusion energy of deuterium.

11.6.1 Origin and Types of Geothermal Energy

Large quantities of heat that are economically extractable tend to be concentrated in places where hot or even molten rock *(magma)* exists at relatively shallow depths in the earth's outermost layer (the crust). Such hot zones generally are near the boundaries of the dozen or so slabs of rigid rock (called *plates*) that form the earth's lithosphere, which is composed of the earth's crust and the uppermost, solid part of the underlying denser, hotter layer (the *mantle*).

According to the theory of plate tectonics, these large, rigid lithospheric plates move relative to one another, at average rates of several centimeters per year, above hotter, mobile mantle material (the asthenosphere). High heat flow also is associated with the earth's hot spots, whose origins are somehow related to the narrowly focused upward flow of extremely hot mantle material from very deep within the earth. Hot spots can occur at plate boundaries (e.g., in Iceland) or in plate interiors thousands of kilometers from the nearest boundary (e.g., the Hawaiian hot spot in the middle of the Pacific plate). Regions of stretched and fault-broken rocks (rift valleys) within plates, like those in

East Africa and along the Rio Grande River in Colorado and New Mexico, also are favorable target areas for high concentrations of the earth's heat at relatively shallow depths.

Zones of high heat flow near plate boundaries are also where most volcanic eruptions and earthquakes occur. The magma that feeds volcanoes originates in the mantle, and considerable heat accompanies the rising magma as it intrudes into volcanoes. Much of this intruding magma remains in the crust, beneath volcanoes, and constitutes an intense, high-temperature geothermal heat source for periods of thousands to millions of years, depending on the depth, volume, and frequency of intrusion. In addition, frequent earthquakes, produced as the tectonic plates grind against each other, fracture rocks, thus allowing water to circulate at depth and to transport heat toward the earth's surface through these fractures. Together the rise of magma from the mantle and the circulation of hot water (hydrothermal convection) maintain the high heat flow that is prevalent along plate boundaries. Accordingly, the plate-boundary zones and hot-spot regions are prime target areas for the discovery and development of high-temperature hydrothermal-convection systems capable of producing steam that can drive turbines to generate electricity (Duffield and Sass 2003).

Water circulates freely through hydrothermal resources. For magma and hot, dry rock resources, much less water is present. The magma resource is at the highest temperature and is relatively water poor. Magmas range in temperature from about 650 to 1300°C, depending on chemical composition. The hot, dry rock resource is characterized by hot, solid rock that contains little or no water because it has few pore spaces or fractures to store and transmit water. The geopressured-geothermal resource consists of hot water saturated with methane and trapped typically in sandstones in young sedimentary basins. In the United States, this resource is abundant along the Gulf Coast in Texas and Louisiana. At present, only hydrothermal resources are used to generate electricity commercially.

11.6.2 Utilization of Geothermal Energy

Geothermal energy has been used for electricity production since 1904 in Italy, and since that time, there has been a steady expansion of this resource on a worldwide basis, with 8000 MW of installed capacity in 2004. High-grade geothermal resources are cost-effective, and in light of their environmentally benign character, the worldwide expansion is expected to continue.

Geothermal energy is immune to daily and seasonal fluctuations and therefore is a reliable power source with a high plant capacity factor. California obtained 5% of its electricity from geothermal plants in 2005. U.S. geothermal plants produce 15 to 16 billion kWh of electricity annually at costs of 4.0 to 8.0 cents per kilowatthour, which is competitive with other energy sources. Seventy plants, totaling 2600 MW of installed capacity, produce electricity in Nevada, Utah, Hawaii, and California. The first resource to be developed in the United States was the steam field at The Geysers, about 90 miles north of San Francisco. The Geysers steam field has produced electricity since 1960 and is the largest geothermal power plant development in the world.

Commercial production of electricity from geothermal resources has been carried out for 40 years in New Zealand and for over 30 years in Japan. Power production is currently under way in Australia, China, Costa Rica, El Salvador, Ethiopia, France (Guadeloupe), Guatemala, Iceland, Indonesia, Italy, Japan, Kenya, Mexico, New Zealand, Nicaragua, Philippines, Portugal (Azores), Romania, Russia, Thailand, and Turkey.

For steam-dominated resources, naturally occurring dry steam can be used directly in a fairly standard, low-pressure steam turbine to generate electricity. The steam is produced via geothermal wells and fed to the turbine through insulated pipes. Power plant developments at The Geysers, California, and Larderello, Italy, use dry steam.

Liquid-dominated geothermal resources are more common than steam-dominated resources. If resource temperatures are sufficiently high (>171°C), the liquid can be flashed partially to steam for use in a steam turbine using a single-flash or double-flash process. If the geothermal reservoir fluid is in a compressed liquid state, it partially flashes to steam in the wellbore as it rises to the surface. Additional steam is separated in surface flash tanks and fed to the turbine. The remaining liquid is injected back into the reservoir. If the resource temperature is high enough, the fluid can be flashed twice. Flashing occurs in the well and in a first separator at the surface. The high-pressure steam is fed to the high-pressure stages of the turbine. The liquid fraction from the first separator is flashed

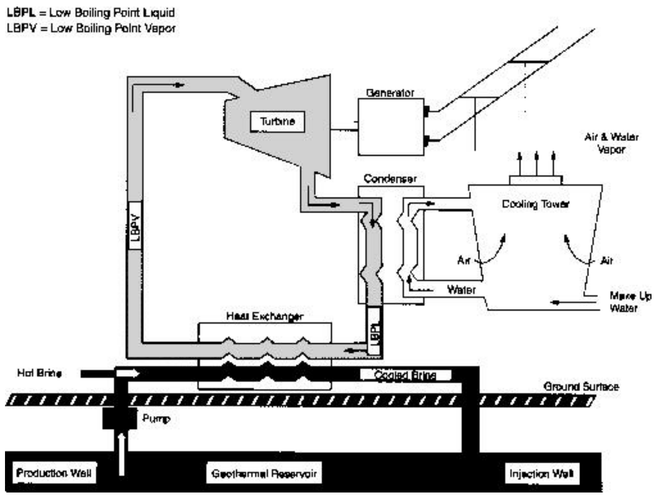


FIGURE 11-13 Schematic diagram of a binary geothermal power plant. (*U.S. Department of Energy*)

again in a second, low-pressure separator. The additional steam is fed to the low-pressure stages of the turbine. The addition of a second flash increases plant efficiency by about 20% compared with single-flash systems.

For liquid-dominated geothermal resources at resource temperatures not sufficiently high enough for flash systems, a binary cycle (organic Rankine cycle) is used (Fig. 11-13). The geothermal fluid is used to heat a working fluid with a low boiling point, such as refrigerants, isobutane, or propane. The heat of the geothermal brine is transferred to the pressurized working fluid in heat exchangers. The working fluid is vaporized and expanded through a turbine to produce power and then condensed and recycled through the heat exchangers in a closed cycle. The cooled brine is injected into the reservoir. Flash and binary cycles are combined sometimes to maximize the efficiency of extracting energy from a single stream of geothermal fluids. Typically, the binary cycle operates in a "bottoming" mode, using the outflow from a flash plant as its input.

In sizing large geothermal power plants, it is important to note that (1) the cost of fluid collection and injection increases as more wells are required, and (2) the relative cost of piping and the amount of heat lost in fluid transmission both increase as the distance between the resource wells and the plant increases. With increased system size, beneficial economies of scale of larger turbine-generators are traded off against the increased cost of the brine-gathering system. Geothermal power costs are sensitive to wellhead temperatures, well flow rates, and well costs. These three parameters determine to a large extent the economic value of a geothermal resource. The thermodynamic efficiency of the conversion process increases with higher resource temperatures. The size and cost of many power plant components vary inversely with temperature and directly with total fluid flow through the plant. Power plant costs increase as resource temperatures decrease because thermodynamic efficiency declines, and the heat content of the geothermal fluid is also less, requiring more mass per kilowatt. At some reservoirs, however, shallow, low-temperature wells can be inexpensive and very productive. Therefore, in some cases, low-temperature geothermal resources still may provide cost-effective power (DOE/EE-0044).

11.6.3 Exploration for Geothermal Energy

Exploration identifies geothermal resources, estimates resource potential, and establishes resource size, depth, and potential production. Refinements of exploration techniques evolved from those used in the petroleum and mining industries. Geothermal exploration of unmapped regions typically proceeds in two basic phases: (1) reconnaissance, or delineation of one or more geothermal provinces; and (2) detailed exploration for exploitable reservoirs within the provinces.

During the first phase, reconnaissance, regional geology and fracture systems are studied, such as young volcanic features, tectonically active fault zones, and overt or subtle geothermal manifestations. To design specific exploration programs, information is collected from various sources, including regional geologic and geophysical maps, satellite imagery, geochemical sampling of thermal and nonthermal waters and gases, analysis of surface rocks and soils, aerial photography, and measurement of thermal gradients in existing wells. If the reconnaissance phase proves satisfactory, the second phase of detailed exploration focuses on individual prospects. The first step in detailed exploration is geologic mapping on scales of about 1:6000. The objective is to understand the broad distribution of rock types, alteration zones, faults, fractures, and thermal features. Geochemical studies also are undertaken, and the collected information is integrated to produce a preliminary model of the area. Geophysical surveys are used to refine the preliminary model. If the model is encouraging, thermal-gradient wells are drilled to depths of 100 to 600 m. These wells provide the first strong direct evidence of the location and intensity of thermal energy. If indications of a hot resource are strong, the first deep hole, a "wildcat well," is drilled. If the wildcat is successful, it is followed by additional production-sized wells to confirm and assess the resource (DOE/EE-0044).

11.6.4 Drilling for Geothermal Energy

Power-generation systems require deep production and injection wells to produce and dispose of the geothermal fluids. The predominant method for drilling these deep wells is rotary drilling, adapted from the oil and gas industry. There are major differences between geothermal well drilling and oil and gas well drilling. Temperatures of geothermal fluids may reach 400°C compared with 200°C for deep oil and gas basins. These high temperatures cause rapid degradation of ordinary drilling equipment and drilling fluids. Production pressures for geothermal reservoirs are usually very low, in many cases subhydrostatic. The underpressured reservoirs can become plugged if ordinary high-density drilling muds are used. To overcome this situation, compressed air, aerated fluids, or foam is used. Geothermal formation rocks are usually more abrasive and harder than petroleum formation rocks. These formations severely degrade bits and tubular materials, especially when air is used as the drilling fluid. High temperatures also cause elongation of the wellhead and casing. Tolerance for casing motions must be accommodated when designing the wellhead, and geothermal wells are cemented from top to bottom to control the effects of casing elongation.

Drilling is one of the most expensive activities in any geothermal power development project. Each well can cost $1 to $3 million, and an average geothermal field consists of 10 to 100 or more wells. Drilling costs account for one-third to one-half of the total cost of a geothermal project.

11.6.5 Geothermal Reservoir Engineering

Reservoir engineering involves (1) formation evaluation, (2) reservoir modeling to forecast productivity and longevity, (3) well field management, and (4) analysis of production trends. Information derived from these activities is used to formulate a production strategy that balances the cost of delivered power or fluid with reservoir longevity.

Formation Evaluation. A number of techniques are used prior to production to evaluate the properties and flow parameters of the reservoir rock, to identify major fluid-bearing fractures and permeable rocks, and to understand reservoir-well interactions. Some of these techniques are analysis of well logs, drill cuttings, and drill cores to derive reservoir parameters such as permeability, porosity, and fracture characteristics; analysis of injected tracer return patterns to define the movement of fluids in the reservoir; pressure-transient well tests to determine the controlling flow and storage capacity of the formation and to delineate reservoir boundaries and heterogeneity; laboratory analyses of reservoir rocks and fluid samples; borehole televiewer, caliper, and spinner (flowmeter) surveys to locate fractures downhole; large-scale seismic arrays to record the pattern of seismic events in the reservoir; magnetotelluric analysis of the reservoir structure; and experimental geophysical and geochemical techniques to locate and map fractures and other geologic features.

Reservoir Modeling. An initial conceptual model of the reservoir, used to estimate its production capacity, is generated in the early stages of production and is updated continuously. The early reservoir model is then expanded from a conceptual model to a complex two- or three-dimensional numerical model. A variety of techniques can be used, including conceptual geologic modeling to define the geometry and physical properties, numerical simulation of reservoir behavior under production and injection conditions, geochemical modeling to analyze changes in reservoir fluids and rocks and to predict the movement of chemical fronts, analysis of well test data to determine key reservoir parameters, and wellbore simulation to analyze fluid flow and heat transfer inside the wellbore.

Well Field Management. Well field management optimizes production strategies. Well field management begins with the initial production stage and continues through the entire production period. Techniques include well field design optimization considering appropriate variations in well locations and depths, production rates and methods, fluid injection locations and rates, and production and injection control strategy; cost-benefit studies of energy extraction rates versus the cost of flashing flow or pumped production; improved injection techniques and injection well replacement strategies to increase production through heat sweep from the reservoir rocks; monitoring subsurface fluid movement using high-precision gravimeters; studies of water-rock interactions, sources of corrosion, and corrosion mitigation; use of liquid-phase and vapor-phase tracers to understand fluid movement within the reservoir; and monitoring microseismic activity to understand the stress vectors in the reservoir and the distribution of productive and/or receptive fractures. Microseismic monitoring also can reduce any environmental concerns that the production or injection of fluids may cause significant seismic activity.

Determination of Reservoir Production Trends. The monitoring of long-term production histories helps to understand reservoir production trends, reservoir–power plant system performance, and reservoir longevity. Recent advances in computer technology facilitate data storage and analysis and enhance the use of reservoir simulators. Methods used are long-term monitoring through geophysical measurements and geochemical sampling to detect fluid depletion and recharge ratios, microseismic monitoring of injected fluid to trace flow patterns and associated fracturing, numerical simulation of nonisothermal flows of multicomponent, multiphase fluids in porous and fractured media, and numerical simulation of multiphase, multicomponent aqueous species and noncondensible gases to model phase partitioning of components, conductive and convective heat flow, relative permeability effects, capillary pressure, and flow in fractured media.

11.6.6 Research and Development

Geothermal research and development activities are continuing worldwide. Improvements in the following areas are expected to have significant impacts: (1) techniques to identify and locate geothermal resources, (2) improved equipment and methods for well drilling and completion, (3) improvements in reservoir engineering, modeling, and management techniques, (4) advanced materials and improved power conversion cycles, and (5) permeability enhancement methods. Exploration techniques being refined include seismic imaging, spontaneous potential, three-dimensional magnetotelluric, gravity, electrical resistance tomography, transient electromagnetic, hyperspectral and satellite imaging, fracture and strain detection, and geochemical surveys. Drilling research and development (R&D) includes high temperature electronics and batteries, polycrystalline diamond compact (PDC) bits, mudjet augmented bits, downhole fiber optics, diagnostics-while-drilling, phosphate and sodium silicate-based well cements, and lost circulation control. For reservoir management, research areas include new chemical tracers, reservoir-wellbore simulators, and injection techniques. Power plant R&D includes innovative power cycles, hybrid wet-dry cooling, noncondensible gas removal, enhanced air-cooled condensers, coatings to prevent corrosion, hydrogen sulfide abatement, and silica recovery from geothermal fluids.

Three hot, dry rock pilot projects are active: the Soultz project in France, the Bad Urach project in Germany and the Cooper Basin project in Australia. These projects involve fracturing rocks of low permeability and conducting water circulation tests to extract heat from the human-made reservoirs. Other hot, dry rock experiments have been conducted at the Fenton Hill site in New Mexico, the Rosemanowes site in Cornwall, England, and the Fjallbacka site in western Sweden.

Techniques to increase permeability in marginally productive natural *hydrothermal* systems are being tried at sites in the western United States. These experiments are being conducted at the margins of productive geothermal fields and are referred to as *enhanced geothermal systems*.

Geothermal heat pump (also called *ground-source heat pump*) technology is making significant strides. The geothermal heat pump uses the earth as a heat source for heating or as a heat sink for cooling. A water and antifreeze mixture circulates through a pipe (usually polyethylene) buried in the ground (vertically or horizontally) and transfers thermal energy to a heat exchanger in the heat pump. The heat exchanger is a water-to-refrigerant loop. The ground-loop heat exchanger rejects heat from the condenser or delivers heat to the evaporator, depending on the mode of operation. Instead of a ground loop, groundwater can be used as the heat source and sink where groundwater is available and disposal of groundwater is acceptable. Geothermal heat pumps offer the advantage of a more stable source/sink temperature when compared with air-source heat pumps and can reduce electricity consumption by up to 30%. Researchers are reducing the cost of ground-loop installations, examining the practicality of adding supplemental heat rejection devices, and developing in situ tests and software to estimate the thermal conductivity of the soil and rock.

Low- to moderate-temperature resources ($<150°C$) are used increasingly for space heating, aquaculture, crop drying, and industrial processes. A well brings hot water to the surface, and a system of piping and heat exchangers delivers the heat to the space or process. A disposal system either injects the cooled water underground or disposes of it on the surface. Worldwide, geothermal direct use installed capacity totals more than 12,000 MWt, with 600 MWt in the United States.

11.6.7 Economics

For 5-MW and larger power plants using high temperature resources, the capital costs for the exploration, wells, and power plant range from $1,150 to $2,100 per kilowatthour. For 5-MW and larger power plants using moderate temperature resources, the capital costs range from $1,350 to $2,500 per kilowatthour. O&M cost range from 0.4 to 0.8 cents per kilowatthour for these plants. These costs are based on World Bank data, and future costs are expected to decline based on technology improvements. DOE/EPRI document EPRI/TR-109496 summarizes additional cost data for flash and binary power plants.

Competition from low-cost electricity from natural gas has strongly affected the development of geothermal power plants in the western United States. Approximately 900 MW of geothermal power plants were installed in the western United States between 1980 and 1990. However, since about 1990, the advent of cheaper electricity from natural gas–fueled systems and low load growth rates have slowed the pace of development. In 1990, geothermal power developers expected to compete against power at 6 to 7 cents per kilowatthour in 1996. By 1993, however, the developers were competing against power at 2.5 to 3.5 cents per kilowatthour in western states. However, strong overseas markets for these systems continue to provide a strong base for ongoing technology improvements and new federd tax credit and . . . increase in the United States are expected to . . . future development.

BIBLIOGRAPHY

Duffield, W. A., Sass, J. H., *Geothermal Energy—Clean Power from the Earth's Heat,*" U.S. Geological Survey Circular 1249, Reston, VA, 2003, http://geopubs.wr.usgs.gov/circular/cl249/.

The World Bank Group, *Geothermal Energy—An Assessment*, http://www.worldbank.org/ html.fqd/energy/geothermal/

Renewable Energy Technology Characterizations, EPRI TR-109496, topical report, prepared by U.S. Department of Energy and Electric Power Research Institute, EPRI, Palo Alto, CA, December 1997.

United States Geothermal Technology: Equipment and Services for Worldwide Applications, U.S. Department of Energy, Assistant Secretary for Energy Efficiency and Renewable Energy, DOE/EE-0044, 1996.

11.7 ENERGY STORAGE

BY MOHAN V. AWARE

Introduction

Energy is manifested in mechanical, electrical, chemical, and thermal forms—the latter being a special form of mechanical energy. Energy may be converted from any of these forms to any other form by one or more intermediate processes. Energy is stored in order to match energy supply with demand and/or to contain it for transport to a point where it can be used. Energy can be stored mechanically, electrically, chemically, or thermally. An energy-storage system generally comprises a converter to alter the energy from the type available to the type best stored, a storage subsystem which contains and stores the energy, and a reconverted subsystem to transform the stored energy to the type needed. The motivation for building and using storage is economics, since storage systems would displace generation equipment fired with premium fuels.

In this section, review of the various types of electrical energy-storage systems that store energy in a variety of forms but whose input and output are electrically presented. The review will include a brief description of the theory underlying the operation, a discussion of engineering concepts including conversion and reconversion methods, and an outline of the specifications and applications for each storage system.

All energy-storage devices have a fixed quantity of usable energy. Once the energy has been consumed, the device will have to be refilled or charged. The quantity of energy or hours of storage of an energy-storage device are determined largely by its intended application and by the incremental cost for the storage subsystem. In the deregulated environment of electrical energy systems, the cost-effectiveness of these storages is having special importance. Research and development is ongoing for all areas of energy storage. Some of the primary energy storage development goals include

- Lower costs
- Longer life
- Higher efficiency

Over the years, a lot of energy storage systems for electrical energy were developed. The most frequently used systems are shown in the Fig. 11-14. When energy is produced in the form of electricity,

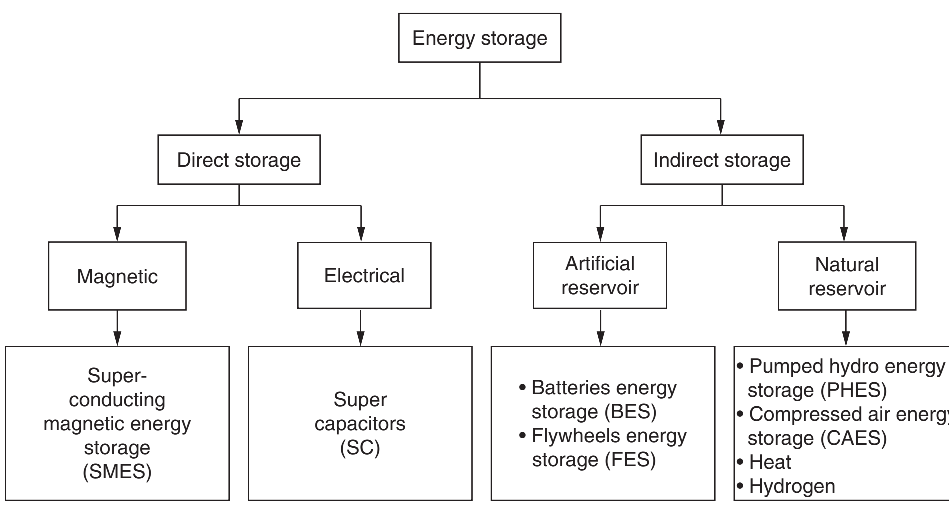


FIGURE 11-14 Energy storage system.

it has to match the load demand. There is no facility to allow the storage of electricity to facilitate the generation interdependently of the load demand. This needs intermittent energy storage devices. The application of these devices in sustainable energy systems is presented with their interface to main system in following sections.

11.7.1 Electrochemical Energy Storage

Batteries. The most traditional of all energy storage for power system is the electrochemical energy storage (EES). The major classification can be made in three categories. They are primary batteries, secondary batteries, and fuel cells. The common feature of these devices is primarily that stored chemical energy is converted into electrical energy. Although the input and output energy of a battery is electrical, the storage is in the chemical form. The most widely known device is the conventional lead-acid battery. The technology is mature and the costs are fairly well-known. The battery life is fairly long, perhaps 10 years or more when properly maintained, but the initial cost is high if one wishes to store enough energy for significant use. In some cases, however, it is the only thing available. In such circumstances, it may be used to operate few lights, television sets, and similar low-demand devices. More effective use of batteries for energy storage awaits the development of more efficient and less costly batteries.

Electrolysis of water represents another technique for storing energy in chemical form. The stored product of hydrogen can be used as fuel. To minimize storage requirements, it must be stored at relatively high pressures, and even so the energy density is fairly low. Greater energy densities could be achieved by liquefaction, but the mechanical and cryogenic technologies required are expensive. One of the major attractions of this approach is that the basic raw material is distilled water, and it is possible to develop similar devices to run on hydrogen directly. These include simple heaters and internal-combustion engines. The former could be used for space heat, grain drying, cooking, etc., and the latter could be adapted to small range vehicles such as mini-tractors and mini-trucks. In certain rural situations, these devices could be useful and would not require imported fuel.

Lead-Acid Batteries. One of the oldest and best-known devices used to store energy is the lead-acid battery. The technology is based on the reduction of lead dioxide to lead sulfate at the positive electrode and the simultaneous oxidation of the lead to lead sulfate at negative electrode. The electrolyte, sulfuric acid, is consumed and energy is discharged during this process. Energy is stored by reversing these reactions, that is, charging the battery.

The energy stored is proportional to the voltage (2.08 V per cell) and to the amount of lead. Lead has a high molecular weight, is an inherently inefficient chemical for battery energy storage, and is used for carrying current in cell. As a result, the amount of lead required per kilowatthour of storage is 50 to 100 lb, which unfavorably affects both the weight and cost of the storage system. A serious drawback of this battery is therefore its low energy density (i.e., energy stored per unit weight of battery), which is about 10 to 25 Wh/lb depending also on design and operation. Under suitable operating conditions, a well-designed and manufactured lead-acid battery can achieve 2000 cycles and last for 10 years. However, life is substantially sacrificed to achieve higher energy densities required for certain application, such as electric vehicles.

A lead-acid battery system comprises individual cells that have capacity range from a few tenths of a kilowatthour to a few kilowatthours. A large submarine cell can have a capacity of 10 kWh and weigh 1,000 lb. Lead-acid cells are often arranged in series strings to achieve the desired dc voltage for conversion or for the application intended. The number of strings will depend on the energy requirements. Energy conversion is often achieved by solid-state converter/rectifier systems. Lead-acid batteries are used in submarines, forklift trucks, uninterrupted power supplies, electric vehicles, and short-term emergency power systems for several applications including telephones, computers, and nuclear power stations. Lead-acid batteries are also being considered for electric utility application. West Berlin's utility (BEWAG) installed a 8.6 MW/9.3 MWh lead-acid battery system in 1987 to meet a part of its system regulation (maintaining frequency stability) and spinning reserve needs (supplying short-term emergency power).

Nickel-Cadmium and Other Commercial Batteries. The nickel-cadmium battery is the only other commonly employed rechargeable battery besides lead-acid. The chemical reaction involves reduction of nickel oxide to nickel hydroxide and oxidation of cadmium to cadmium hydroxide in an alkaline electrolyte (20% to 35% potassium hydroxide). The voltage of this couple (1.3 V) is less than that of lead-acid, and the energy density (about 20 Wh/lb) is slightly better. Due to the high costs for nickel and cadmium, this battery is expensive and is only used when extremely long life or lighter weight is required. Major applications are satellites, portable equipment and tools, and various military applications. Other secondary (rechargeable) batteries considered for these and their specialized application are silver-zinc, nickel-iron, and nickel-hydrogen. These systems are all expensive and are less environmentally friendly, but they are durable and have very long lives. Compared to lead-acid battery, the higher reliability of individual cell is offset by the larger number of cells required.

Advanced Batteries. Over the past two decades, a number of advanced batteries have been investigated. The aim of the research programs has been to lower the costs while maintaining the desirable specifications (durability and performance) of the lead-acid and nickel-cadmium batteries. A secondary objective is to achieve improved energy densities. The advanced battery systems closest to commercialization are zinc-chloride, zinc-bromide, and beta (sodium-sulfur) batteries. These are compared in Table 11-7. In contrast to the lead-acid and nickel-cadmium systems, these advanced batteries use low-cost, readily available active materials and enjoy simple electromechanical reactions, which should lead to excellent durability.

These advanced batteries either operate at higher temperature (Na-S) or employ flowing electrolytes (Zinc-halogen). The additional subsystems for such batteries to maintain flow or temperature add to overall system complexity, which in turn, affects their reliability. However, these subsystems normally will be built within the battery system and will have the same modular character of the lead-acid battery. As a result of these additional subsystems, optimal economics will not occur until module sizes are in the 100- to 400-kWh range, a factor of 10 larger than the largest lead-acid battery modules. The size of module will have application only where there are fairly sizable energy storage requirements, such as electric utility and commercial or industrial applications, which require a capacity of few hundred kilowatts to several megawatts. While there is hope that these advanced batteries can be used for electric-vehicle application, their economic optimal size and complexity suggest that the economic goals for this application will be difficult to achieve. An electric-vehicle battery has to be inherently simple from both the design and engineering standpoint to meet owner expectations.

The advanced batteries are not yet commercially available at affordable costs. Engineering prototypes of the zinc-chloride system at a size of 500 kWh began independent testing in early 1984. Similar-sized prototypes of the other advanced batteries were tested in 1991. Limited commercial availability of these systems can occur within a few years of successful independent testing. However, acceptable costs may occur for yet an additional 3 to 5 years, and only with a reasonable market.

Fuel Cells. A *fuel cell* is an electrochemical device wherein the chemical energy of a fuel is converted directly into electric power. The main difference between a conventional battery and a fuel cell is that, unlike a battery, a fuel cell is supplied with external reactants; as a result whereas a battery is discharged, a fuel cell never faces such a problem as long as supply of fuel is provided. Fuel cells, distinguished from other secondary batteries by their external fuel store, have an even longer history than the lead-acid battery. The first hydrogen-oxygen fuel cell was demonstrated in principle by English lawyer W. R. Grove in 1989, although the bulk of fuel cell's development has been in the

TABLE 11-7 Advanced Battery Systems

System	Positive electrode	Negative electrode	Cell voltage	Temperature, °C	Achievable energy densities, Wh/lb
Na–S	S	Na	2.1	300–350	45–55
Zn–Cl	Cl	Zn	2.1	20–50	30–40
Zn–Br	Br	Zn	1.8–1.9	20–50	20–30

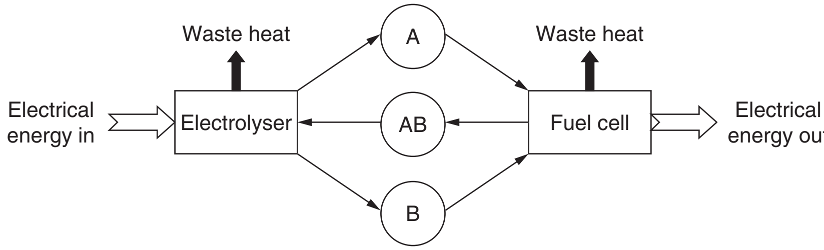


FIGURE 11-15 Schematic diagram of an electrochemical fuel cell.

last 40 years and the major application has been in the space industry. Major research activity in secondary batteries worldwide is concerned with material research for advanced batteries, in particular, materials for solid solution electrodes and for solid electrolytes. It should be noted that solid electrolytes are equally applicable to an electrochemical fuel cells. The processing functions of an electrochemical fuel cell are presented in Fig. 11-15.

Essential functions of the fuel cell are

1. The charging (or electrolyser) function in which the chemical AB is electrochemically decomposed to A and B.
2. The storage function in which A and B are held apart.
3. The discharge (or fuel cell) function in which A and B are reunited with simultaneous generation of electricity.

Fuel cells are generally characterized by the type of electrolyte that they use. Main fuel cell systems under development for practical applications are phosphoric acid (PA), proton exchange membrane (PEM), molten carbonate (MC), solid oxide (SO), direct methanol (DM), and alkaline fuel cell.

The different types of the fuel cells are having their respective properties and applications. A typical fuel cell-based power processing system showing the major plant processes is shown in the Fig. 11-16 . There are three major steps involved in the generation of the power from a fuel cell. The first and foremost step is achieved purity of the available hydrogen gas. This is done with the help of a fuel processor. A carbonaceous fuel is fed to the fuel processor, which in turn, produces a hydrogen rich gas at its output. This hydrogen rich gas is then fed to the anode electrode of the fuel cell. The second step involves the fuel cell operation itself. The fuel cell is fed with the hydrogen rich gas at its anode and supply of air to the cathode. The hydrogen atoms at the anode gets split up into positive protons and negative electrons. These electrons follow an external path on their way to the cathode, thus supplying power to an external load in the process.

The third step is the power conditioning step, which includes power electronic converters. Power electronic converters add more flexibility to the operation of the system.

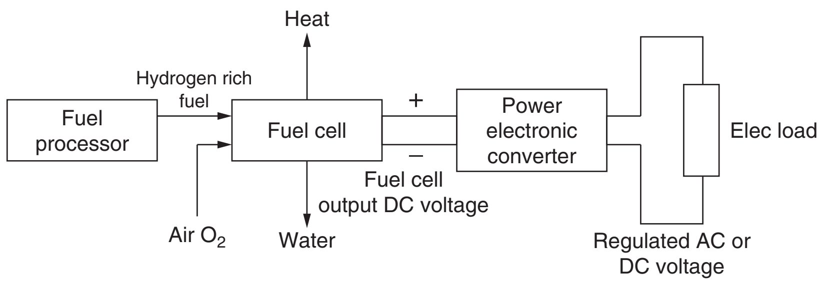


FIGURE 11-16 Typical fuel cell-based power processing system.

11.7.2 Mechanical energy storages

Flywheels. Mechanical energy may be stored in the form of kinetic or potential energy. A flywheel, in essence, is a mechanical battery—simply mass rotating about an axis. Although the use of the flywheel as a kinetic energy storage device is well-known, it has not, until recently, been considered practical to build flywheels that could store energy for several hours or days. They take an electrical input to accelerate the rotor up to speed by using the built-in motor, and return the electrical energy by using this same motor as a generator. Steel flywheels have been used to power certain types of vehicles, such as buses and trucks, for very short runs, but the energy-storage periods have at the best been for a few minutes.

The basic problems of flywheels are as follows. First, the stress in any rotating structure is proportional to the first power of the diameter and second power of the angular velocity. With steel structures, the limiting stresses are exceeded before any significant amount of usable energy can be stored. Second, even if the flywheel could store significant energy, it would soon be dissipated by bearing friction and windage losses. It has been pointed out by Post and Rabenhorst that one or two orders of magnitude improvement over conventional flywheel technology could be achieved by

(*a*) Using fiberglass resins and polymer materials, which have a much higher strength-weight ratio than steel.

(*b*) Running the flywheel in a vacuum.

(*c*) Using air-or magnetic-suspension bearing technology developed in recent years.

If the projections of Post and Rabenhorst are supported by experimental evidence, it appears that flywheel could become practical for storing significant amount of energy efficiently for several hours. This could make flywheel practical for electrical peaking and short-run vehicular transportation.

Flywheels involve the storage of kinetic energy in a rotating object. The stored energy is proportional to the object's moment of inertia times the square of its angular velocity. In order to optimize the energy-to-mass ratio, the flywheel needs to spin at the maximum possible speed. For a flat disk, the moment of inertia is proportional to its mass times the square of its radius. Therefore, the energy stored is proportional to its mass and the square of angular velocity. Rapidly rotating objects are subject to centrifugal forces that can rip them apart. Thus, while dense material can store more energy, it is also subject to higher centrifugal force and thus fails at lower rotational speeds than low-density materials. Therefore, tensile strength is more important than the density of the material. Two approaches to flywheel energy storage are generally pursued. The commonly used approach is to use a heavy material, like steel, and operate at a moderate speed. The other approach, the so-called advanced flywheel, uses lighter high-strength materials like glass or carbon fibers and operates at high speed. Recent advances in composite materials technology may allow nearly an order of magnitude advantage in specific strength of composite when compared to even the best engineering metals. The result of this continuous research in composite has been a flywheel that operates at rotational speed in excess of 100,000 rpm with tip speeds in excess of 1000 m/s.

Flywheels have smoothed our energy sources in such historic applications as the potter's wheel, the grain mill, and the water wheel. Today, its best use is in engines of various types. For large-scale electric storage, economics would dictate the use of lighter-weight high-speed (30,000 r/min) wheels operating in a vacuum and using magnetic bearings to reduce friction. Under these conditions, energy efficiencies of about 80% are achievable. Since the speed of the wheel will change as energy is added or withdrawn, the motor-generator will have to accommodate variable speeds. Several approaches to variable-speed generation have been investigated for wind application, although today's wind machines use fixed speed generators. Conservation measures have encouraged the use of variable-speed motors, and several manufactures now have this technology available in large sizes, for such applications as thermal power plant fans and pumps. The flywheel itself is an expensive way to store energy, with costs probably twice that of the lead-acid battery. However, it has several advantages over chemical energy storage, like energy density and better capability to exchange energy.

However, for application such as utility system regulation, where there are high-power and low-energy requirements, flywheels would be suitable if integrity and reliability can be demonstrated. The ideal use for today's flywheels is still to absorb and dissipate mechanical energy to smooth the

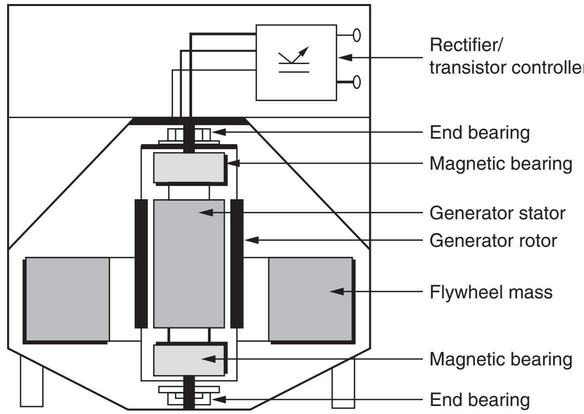


FIGURE 11-17 Flywheel storage system.

operation of rotating machinery. The ultrahigh rotational speed required to store significant kinetic energy in these systems virtually rules out the use of conventional mechanical bearings. Instead most systems run on magnetic bearings. These relatively recent innovations use magnetic forces to leviate a rotor, eliminating the frictional losses inherent in rolling element and fluid film bearing. Unfortunately, aerodynamic drag losses force most high speed flywheels to operate in a partial vacuum. A modern high-speed flywheel system is shown in Fig. 11-17.

Pumped Storages. Mechanical energy may be stored in the form of potential energy by pumping water to a reservoir at a higher level and allowing it to flow down to a lower reservoir whenever the energy is required. In this pump-back system, the kinetic energy of the flowing water is usually converted to electrical energy, although it can be used directly as mechanical energy. The energy can be extracted from the hydro power plant depending on both the volume of the water available and the head of the water that can be exploited. The pump storage will provide the most efficient and cheapest operation when it can be provided a high head between the two reservoirs. This will allow greatest amount of energy to be stored in the smallest volume of water resulting in a smaller pump and turbine, reducing capital costs. Pumped hydro storage usually comprises the upper reservoir, waterways, a pump, a turbine, a motor, a generator, and lower reservoir as shown schematically in the Fig. 11-18.

A difficulty of the pump-back storage system is that relatively large reservoirs are required. Most pump-back systems are built where there is a large natural basin for an upper reservoir located near an existing lake or river, the latter serving as the lower reservoir. With current technology in pump-back systems, generally about two-thirds of the energy is available for utilization after storage.

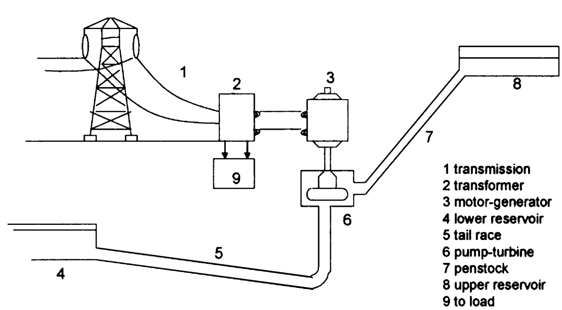


FIGURE 11-18 Pumped hydroelectric energy storage.

Hydraulic turbines have a maximum efficiency of around 95% and pumps are less efficient operating around 90%. Therefore, pump storage plants will have maximum efficiency of around 85%.

Pumped hydro or pumped storage, as it is often called, has been used in electric utility systems for over 50 years. Electricity is used to operate a motor-generator combination to pump water to an elevated reservoir. When energy is required, water is allowed to flow down to a lower reservoir through a turbine-generator combination, much like the turbine generators in conventional hydroelectric plants. The energy stored is proportional to the head (height differential between the upper and lower reservoir) times the stored volume of water. For head of 1200 ft, 36 ft of water is required for generation of 1 kWh. For a head of 120 ft, 360 ft of water is required to generate 1 kWh.

In areas where topographic or ecological considerations rule out pumped-hydro plants, utilities now may choose to go underground. A modification of pumped hydro, underground pumped hydro (UPH) would have its lower reservoir located about a mile below the earth's surface in a competent hard-rock cavern. A UPH plant therefore can be sited in flat terrain and should have only a minimal effect on the surrounding environment. The most significant drawback of UPH is that the plants must be in huge sizes (2,000 MW/20,000 kWh) to be economically attractive.

Both pumped hydro and UPH, like conventional hydro, have very fast response characteristics (emergency full-power capability is 10 s) and very high efficiencies (72% to 75%). These factors, along with low capital costs, have led to the construction and operation of 18,000 MW of pumped hydro in the United States. Another 16,000 MW is planned or under construction. Pumped hydro typically will be the storage alternative of choice for electric utilities having favorable topography. Main barriers for the future growth in pumped-hydro capacity are the availability of suitable sites and large environmental impact of these schemes.

Compressed Air. Another method for storing mechanical energy is through compressed air. Two basic methods are of interest: direct compression and isothermal compression. Direct compression, to be practical, requires the availability of natural or man-made underground caverns, which can be sealed off and pressurized to hydrostatic pressure or below. Utilization of the stored energy requires the availability of low-pressure turbines or piston engines similar to the old-style steam engines. Generally, such systems provide a relatively low overall efficiency. About one-third of initial energy can finally be utilized. This figure can be improved considerably when compressed-air system is combined with gas turbine cycles or with refrigeration cycles, because the energy required to compress the turbine inlet air need not be charged against turbine fuel efficiency if the air is already compressed.

Of increasing interest is isothermal hydraulic compression. In this process (which is at least 3,000 years old), a vertical tube is placed in front of a dam which is creating a head of water (typically 5 to 10 m). Water flowing into the inlet entraps air bubbles, which are carried with water. The water then passes through the bottom of a sealed tank or cavern, where the air bubbles rise. Ultimately, the air in the tank or cavern is compressed to a pressure slightly under the hydraulic head of the vertical (venturi) tube. That compressed air can then be used in a simple turbine or other suitable expander to obtain usable mechanical energy. Such isothermal hydraulic compressors were used from ancient times until the later part of the nineteenth century, when they were gradually replaced by electrical, steam, or internal-combustion devices. Because the compression is isothermal, the apparent efficiency of such systems can be very high, and with the rising cost of fuel, interest in this storage system is rapidly rising. A typical air-storage gas-turbine power plant with a constant pressure reservoir is shown in Fig. 11-19.

CAES involves the compression of air into an underground cavern or reservoir (excavated rock, solution-mined salt, or aquifer) where it is stored. Favorable geologies exist in about 75% of the United States. Geologic considerations generally dedicate that most of the heat of compression be removed (and used or stored, if possible) before the air is stored underground. During generation, the air is expanded through a turbine to drive a generator. Economics dictates that heat be added to the air from a thermal store or through combustion of fuel prior to the expansion-generation phase.

CAES technology recently has become commercialized for electric utility application. The commercial units incorporate modified combustion turbines wherein the compressor and expander are physically separated by clutches and motor-generator. For the expansion-generation phase, the technology uses oils or gas, although other fuels can be used. When a heat exchanger is incorporated to use exhaust heat to raise the temperature of the pressurized air, the overall energy balance involves approximately

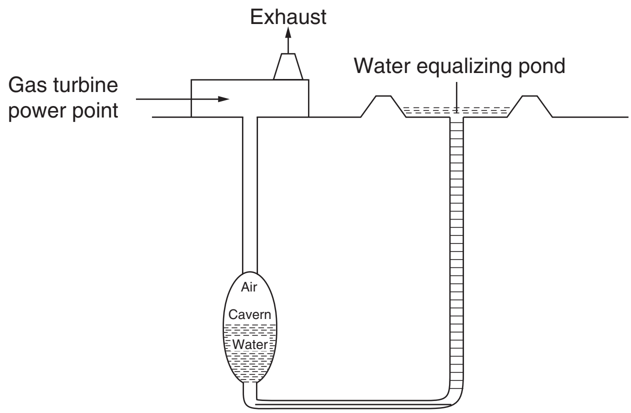


FIGURE 11-19 Air-storage gas-turbine power plant with a constant pressure reservoir.

4,000 Btu of oil or gas (during the generation phase) and 0.775 kWh of electricity (for the compression step) to produce 1 kWh of plant output. Typical energy-storage efficiencies cannot be calculated because a mixture of fuels is used. However, overall energy use is about 11,500 Btu/kWh, assuming the electricity for compression was generated at 10,000 Btu/kWh. A storage device using only electricity would have an efficiency of 87% to achieve an overall fuel use of 11,500 Btu/kWh, again assuming charging electricity is generated at 10,000 Btu/kWh. CAES system can be used on very large scales. Unlike other systems considered, large-scale CAES is ready to be used with entire power plants. Apart from the hydro pump, no other storage method has a storage capacity as high as CAES. Typical capacities for a CAES system are around 50 to 300 MW. The storage period is also the longest due to the fact that its losses are very small. A CAES system can be used to store energy for more than a year. The main drawback of CAES is probably the geological structured reliance.

11.7.3 Thermal Energy Storage

Energy can be stored as heat in appropriate materials, provided that suitable thermal insulation surrounds the storage substances. Hot rocks and fireplace bricks have served as primitive heat storage devices from ancient times. Water, rocks, and materials of high heat capacity, which are generally available, are commonly used as a storage media. Sources of energy such as solar heaters or wind generator can be used. The major difficulty is that the energy can be reutilized only as heat. Reconverting the heat energy to electrical or mechanical energy involves significant losses because of fundamental thermodynamics limitations. Storage temperature generally would be low, of course, and maximum reconversion efficiency (to mechanical or electrical form) would be on the order of few percent. Nevertheless, direct reutilization as heat can be practical, and this is done in some places. While thermal and cool storage has many applications, including thermal storage of off-peak energy in many European homes, it is not economically viable with electricity being both the energy source and the final product. The poor efficiency of converting heat to electricity is prohibitive. Thermal storage is generally only economically attractive when the energy required is heat. Thermal energy storage is ideally suited for applications such as space heating, where low quality, low temperature energy is required, but it is also possible to use thermal energy storage with conventional coal and nuclear-fired power plants, which dominates installed capacity of electricity utilities and is likely to continue to do so for the near future. Site-specific applications where heat is stored as steam and hot water at power plants have proven economic. However, retrofits of such technology into existing power plants have proven economic, but it would be difficult and costly since thermal and cool storage do not involve electricity and effectively store the thermal energy. Small water tanks are widely used for solar heat storage systems as shown in the Fig. 11-20.

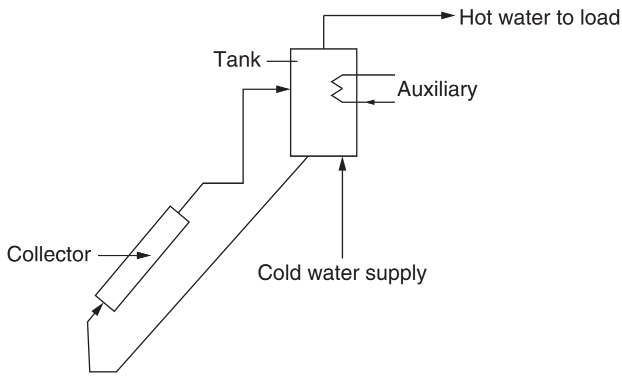


FIGURE 11-20 Hot water panel system with natural circulation.

11.7.4 Electrical Energy Storage

Superconducting Magnetic Energy Storage. Electricity (dc) can be stored in magnetic fields, wherein the energy stored is proportional to the inductance times the square of the current. An energy storage system based on magnetic would involve an ac-dc-ac converter system (like batteries) and a large coil of a superconductor (e.g., a niobium-titanium alloy). To achieve superconductivity, the conductor is maintained in a bath of liquid helium at about 1.8 K. Since there are no moving parts in the coil and electrical resistance is near zero, efficiencies can be above 90%. Energy losses occur in the converter and refrigeration system, although both can be designed to achieve high efficiency. The cost per unit of stored energy of superconducting magnetic energy storage (SMES) varies with the minus one-third power of stored energy. In other words, cost per unit of stored energy decreases by 21% for each doubling of storage capacity. In practice, therefore, very large sizes of several thousand megawatthours are required for SMES to be cost-competitive.

The basic operation of a complete SMES system is very simple and its setup is shown in Fig. 11-21. The transmission voltage from ac network is first stepped down from a few kilowatts to several hundred volts using a step-down transformer. This is then converted into dc, which is fed into the superconducting coil.

When power flows from system to coil, the dc voltage will charge up the superconducting coil and energy is stored in the coil. When ac network requires power boost, say when there are sags, spikes, and voltage and frequency instabilities, coil discharges and acts as a source of energy. The dc voltage is converted back into ac voltage through the converter.

The maximum energy stored depends on the design of the device. SMES systems are able to store upto 10 MW. However, some research believe an SMES can potentially store up to 2,000 MW. Theoretically, a coil of around 150 to 500 m radius would be able to support a load of 5000 MWh at 100 MW depending on the peak field and ratio of the coil's height and diameter.

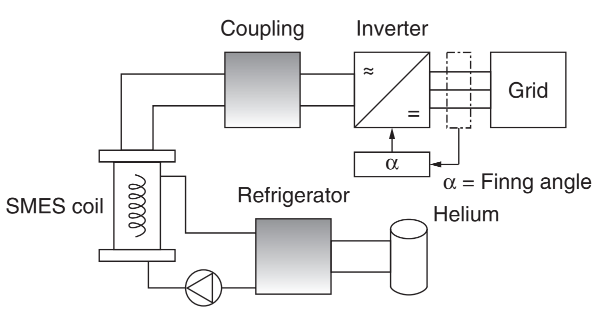


FIGURE 11-21 Basic setup of a SMES unit.

SMES is, however, attractive for applications where high power is required for very short periods. For example, a small 10- MW, 3-s experimental system was built to inhibit subsynchronous resonance in a high-voltage ac line from the state of Washington to California. Through R&D, SMES costs might be reduced sufficiently to become competitive for bulk-storage application. Ongoing research is directed to improving conductor current density and developing innovative design concepts that lower conductor support requirements. High Temperature Superconductor (HTS) coils of Ag-clad $(Bi, Pb)_2 Sr_2 Ca_2 Cu_3 O_{10+x}$ material, which can be operated with liquid nitrogen are commercially available. With successful research, issues involving the effects of magnetic fields, land use, and quality control suggest that the practical systems are commercially available and operational. A typical bidirectional converter with cryogenic arrangement is shown in Fig. 11-22.

Ultracapacitor. Capacitors are some of the most essential building blocks of electronic circuits to hold dc voltages. Based on the same principle, but on a much larger scale, it is conceivable that capacitors could be used to store energy for extended periods of time. Until some time ago, however, capacitors only managed to hold very little energy compared to a regular battery. In 1997, researchers from CSIRO developed the first supercapacitor. This is basically a capacitor which is able to hold significantly more charge using thin film polymers for the dielectric layer. The electrodes are made of carbon nanotubes. The energy density of a normal capacitor is only 0.5 Wh/kg. PET super capacitors can store 4 times more energy compared to the normal capacitor. Carbon nanotubes and polymers are practical for supercapacitors. Carbon nanotubes have excellent nanoporosity properties allowing the polymer tiny spaces to sit in the tube and act as a dielectric. Polymers have redox (reduction-oxidation) storage mechanism along with a high surface area. There are researches going on to replace carbon nanotubes with ceramics for their superconducting properties. These promising technologies introduce potential to improve the energy storage and represent a new generation of electrochemical components with very high capacitance for energy storage. A capacitor storage system with grid interface is shown in Fig. 11-23.

Supercapacitors are well suited to replace batteries in many applications. This is because at the moment their scale is comparable to that of batteries, from small ones used in cellular phones to large ones that can be found in cars. Even though supercapacitors have a lower energy density compared to batteries, they avoid many of the battery's disadvantages.

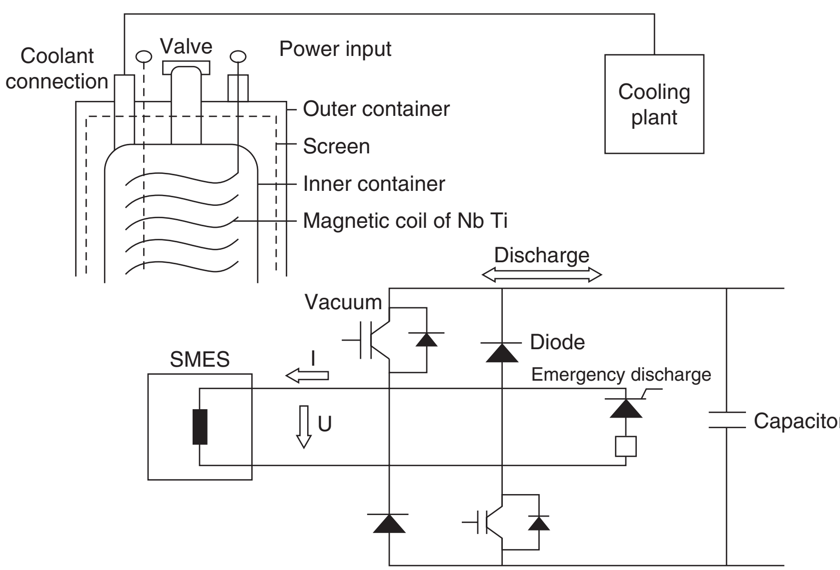


FIGURE 11-22 SMES high-temperatures system (HTS).

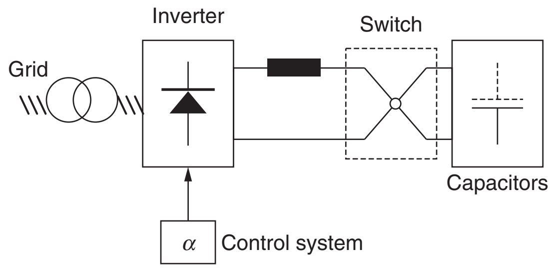


FIGURE 11-23 Capacitor storage system with grid connectivity.

Batteries have a limited number of charge/discharge cycles and take time to charge and discharge because the process involves chemical reactions with noninstantaneous rates. These chemical reactions have parasitic thermal release that causes the battery to heat up. Batteries have a limited life cycle with a degrading performance and acidic batteries are hazardous to the environment. Super capacitors can be charged and discharged almost an unlimited number of times. They can discharge in matters of milliseconds and are capable of producing enormous currents. Hence, they are very useful in load levelling applications and fields where a sudden boost of power is needed in a fraction of a second. They do not release any thermal heat during discharge. Supercapacitors have a very long lifetime, which reduces maintenance costs. They do not release any hazardous substances that can damage the environment. Their performance does not degrade with time. Supercapacitors are extremely safe for storage as they are easily discharged. They have low internal resistances, even if many of them are coupled together. Even though they have a lower energy density, are bulkier and heavier than an equivalent battery, they have already replaced batteries in many applications due to their readiness in releasing power. Supercapacitors were initially used by the U. S. military to start the engines of tanks and submarines. Most applications nowadays are in the field of hybrid vehicles and handheld electronic devices.

In most hybrid vehicles, 42 V supercapacitors are used. General Motors has developed a pickup truck with a V8 engine that uses the supercapacitor to replace the battery. The efficiency of the engine rose by 14%. The supercapacitor supplies energy to the alternator. In rural areas, where there are voltage sags in the power grid, supercapacitors can be used to reduce the effect of fluctuations. The supercapacitor has become available to the public. A commercial supercapacitor can hold 2,500 F, release 300 A of peak current with a peak voltage handling of about 400 V. The life cycle of this supercapacitor is more than 1×10^6 charge/recharge cycles.

11.7.5 Economics of the Energy Storage Media

There is a diversity of storage technologies to store electrical energy. They are somewhat more complex to analyze than that of conventional generation systems because the number of hours of operation will dictate capital costs and its use. It is important to note that the total cost of storage system ($/kW) is the sum of power-related costs expressed in $/kW and the storage-related costs ($/kW) times the required hours of storage. Thus, if long periods of storage are required, it is critical to have an inexpensive storage subsystem. On the other hand, for high-power applications with short discharge periods, low conversion cost and fast response are the key requirements. Table 11-8 outlines relative economics for the various storage systems discussed. For electric utility systems, a combination of storage systems might be optimal because of the relative costs. For example, batteries could be used for peaking (less than 5 h of storage) and CAES for intermediate duty cycle (more than 8 h of storage).

Besides economics, there are several other considerations that should be addressed before a storage system is selected for a particular application: siting restrictions, environmental impact, response time, unit size, lead time, commercial availability, and operational experience. A relative comparison of specific energy and specific power for various storages and their backup times are shown in Fig. 11-24 *a* and *b,* respectively.

TABLE 11-8 Relative Economics for Energy Storage System

Storage technology	Conversion cost, $/kW	Storage cost, $/kW
Lead-acid battery	Low	High
Nickel-cadmium battery	Low	Very high
Advanced battery	Low	Moderate
Pumped hydro	Moderate	Low
Underground pumped hydro	Moderate	Low to moderate
Compressed air	Moderate	Low
Superconducting magnet	Low	Very high
Flywheel	Low to moderate	Very high

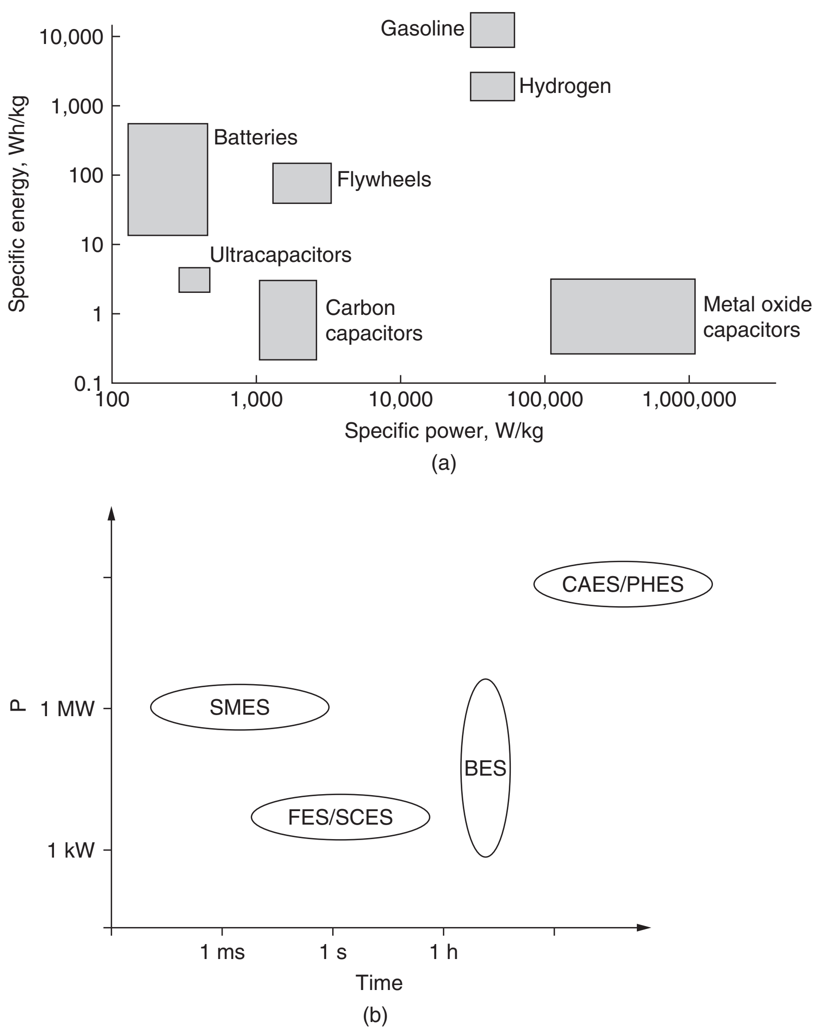


FIGURE 11-24 Electrical power and energy storage comparison: (a) comparison of specific energy and specific power in various energy storages; (*b*) storage technology with its backup times.

TABLE 11-9 Performance and Cost Comparison of Energy Storage

	Batteries	Flywheel	Pumped hydro	CAES	GAS	Micro SMES	SMES	Capacitors
Efficiency	~75%	~90%	~75%	~70% + fuel	~60% + fuel	~95% + refrigeration	~95% + refrigeration	~95%
Energy Capital Cost	200 $/kWh + replacement	800 $/kWh	16.8 $/kWh	10 $/kWh	250 $/kWh	~300 $/kJ	Depends on size	3600 $/kJ
Power Capital Cost $/kW	300	220	600	425	1000	500	300	300
Fixed O & M $/kW/yr	1.55	7.5	4.3	1.35	2.8	8	1	5% of capital
Variable O & M $/kWh	0.5	0.4	0.43	0.1	0.1	0.5	0.1	-

Only after these entire factors are considered then the storage system be selected for any particular application. Once selected, storage systems often will be very competitive with generation technologies, provided low-cost charging energy is available.

Table 11-9 presents the most commonly used and emerging storage systems' comparison with their operating hours and efficiencies. Most storage systems have greater size and operational flexibility than the competing generation schemes and therefore should be seriously considered in any utility expansion plan and for certain utility customers.

GLOSSARY

BES-Battery energy Storage
CAES-Compressed Air Energy Storage
EES-Electrochemical Energy Storage
FES-Flywheel Energy Storage
PHES-Pumped Hydropower energy storage
SMES-Superconducting Magnetic Energy Storage

BIBLIOGRAPHY

Sels, T., Dragu, C., Craenenbroeck, T., and Belmans, R., "New Energy Storage Devices for an Improved Load Managing on Distribution Level," Proc. *IEEE Porto Power Tech. Conference*, Sept. 10–13, 2001, pp. 6.

Schoenung, S. M., and Burns, C., "Utility Energy Storage Applications Study," *IEEE Trans. on Energy Conversion*, vol. 11, issue 3, Sept. 1996, pp. 658–665.

Buckles, W., and Hassenzahl, W. V., "Superconducting Magnetic Energy Storage," *IEEE Power Engineering Review*, vol. 20, No. 5, May 2000, pp. 216–220.

Hassenzahl, W. V., "Superconductivity, an Enabling Technology for 21st Century Power Systems." *IEEE Trans. on Applied Superconductivity*, vol. 11, No. 1, March 2001, pp. 1447–1453.

Bullard, G. L., Sierra-Aleazar, H. B., Lee, H. L., and Morris, J. L., "Operating Principals of the Ultracapacitor," *IEEE Trans. in Magnetics*, vol. 25, no. 1, Jan.1989 pp. 102.

Ter-Gazarian, A., Energy Storage for Power Systems, IEE Series 6, Peter Peregrinus, London, 1994.

Emadi, Ali, Ehsani, Mehrdad, Miller, J. M., Vehicular Electric Power Systems, Marcel Dekker, Inc., New York, 2004.

Energy Storage Association, Technologies and Applications. Available at http://www.energystorage.org/ technologies.htm

American Superconductors, Inc., Technical Papers. Available at http://www.amsuper.com/semsfact.htm

11.8 BATTERIES

BY ROBERT D. WEAVER AND PAUL BUTLER

11.8.1 Principles of Operation

Electrochemical Principles and Reactions. A *battery* is a device that converts the chemical energy contained in its active materials directly into electrical energy by means of oxidation-reduction electrochemical reactions. These types of reactions involve the transfer of electrons from one material to another. In a battery (Fig.11-25), the negative electrode or anode is the component capable of giving up electrons, that is, being oxidized during the reaction. It is separated from the positive electrode or cathode, that is, the component capable of accepting electrons and being reduced. The transfer of electrons takes place in the external electric circuit, connecting the two materials. Transfer of charge is completed within the electrolyte by movement of ions (anions and cations), not by electron flow.

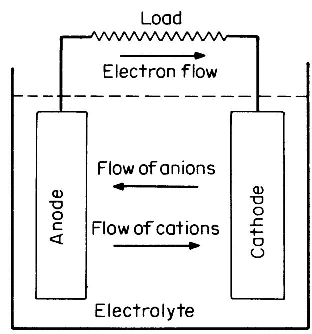


FIGURE 11-25 Electrochemical operation of a battery.

The operation of a fuel cell is similar to that of a battery except that one or both of the reactants are not contained in the electrochemical cell but are fed into it from an external supply when power is desired. The fuels are usually gaseous or liquid (compared with the metal anodes generally used in batteries), and oxygen or air is the oxidant.

Components of Batteries. The basic unit of the battery is the cell. A battery consists of one or more cells, connected in series or parallel depending on the desired output voltage and capacity. The cell consists of three major components: the anode (the reducing agent or fuel), the cathode or oxidizing agent, and the electrolyte which provides the necessary internal ionic conductivity. Electrolytes are usually liquid, but some batteries employ solid electrolytes which are ionic conductors at battery operating temperatures. In addition, practical cell design requires a separator material (which serves to separate the anode and cathode electrodes mechanically), electrically conducting grid structures or materials added to each electrode to reduce internal resistance, and suitable containers.

Theoretical Cell Voltage and Capacity. The theoretical capacity (amperehours) of a battery system is determined by its active materials. The maximum electrical energy (watthours) corresponds to the free-energy change of the reaction. The theoretical voltage and specific energy ratings of a number of electrochemical systems are given in Table 11-10. The voltage is determined by the active materials selected, while the amperehour capacity is determined by the amount (weight) of available reactants. One gram-equivalent weight of material will supply 96,480 c, or 26.805 Ah, of electrical charge.

Factors Influencing Battery Voltage and Capacity. In practice, only a small fraction of a battery's theoretical specific energy is realized. As a rule of thumb, a secondary (rechargeable) battery will not attain much more than 25% of the theoretical value; claims of specific energy values significantly greater than 25% of theoretical should be viewed with appropriate caution. Refer to Table 11-10

TABLE 11-10 Characteristics of Some Battery Systems

Battery type—common name	Chemical reactant		Theoretical battery		Practical battery	
	Negative electrode	Positive electrode	Voltage, V	Specific Energy, Wh/kg	Nominal voltage, V	Nominal specific energy,[a] Wh/kg
Primary						
Leclanche	Zn	MnO_2	1.5	260	1.2	80
Magnesium	Mg	MnO_2	2.0	582	1.5	125
Alkaline MnO_2	Zn	MnO_2	1.6	380	1.3	95
Mercury-zinc	Zn	HgO	1.34	260	1.2	95
Mercad	Cd	HgO	0.9	160	0.85	45
Silver-zinc	Zn	AgO	1.7	480	1.6	130
Li–MnO_2	Li	MnO_2	3.5	1000	2.7	200
Li–sulfur dioxide	Li	SO_2	2.95	1100	2.9	250
Li–thionyl chloride	Li	$SOCl_2$	3.66	1500	3.5	340
Zinc-air[b]	Zn	Air (O_2)	1.6	1310	1.2	200
Hydrogen-oxygen fuel cell	H_2	O_2	1.23	3660	1.0	100
Secondary						
Lead-acid	Pb	PbO_2	2.1	175	1.8	40
Edison	Fe	NiOOH	1.5	230	1.2	40
Nickel-cadmium	Cd	NiOOH	1.3	210	1.2	50
Silver-zinc	Zn	AgO	1.85	440	1.5	140
Lithium ion[i]	Li (C6)	Li0.5CoO_2	3.9	390	2.3	110
Lithium polymer	Li	V6O13	2.8	880	2.3	150
Nickel-zinc	Zn	NiOOH	1.75	330	1.6	70
Zinc-air[b]	Zn	Air (O_2)	1.6	1030	1.1	150
Nickel–metal hydride	H_2 · (metals)	NiOOH	1.3	230	1.2	60
Zinc-bromine	Zn	Br_2	1.8	430	1.0	75

Reserve						
Sea-water[c]	Mg	CuCl	2.5	600	1.6	80
Silver-zinc[d]	Zn	AgO	1.85	480	1.5	30
Thermal[e]	Li Alloy	FeS_2	1.9	500	1.6	25
High temperature						
Sodium-sulfur[f]	Na	S	2.07	750	1.9	100
Zebra[g]	Na	$NiCl_2$	2.6	790	2.4	100
Lithium · (aluminum)| iron disulfide[h]	Li · Al	FeS_2	1.8	640	1.4	100
Lithium · (aluminum)| iron sulfide[i]	Li · Al	FeS	1.3	450	1.2	80

[a]Delivered energy when discharged at normal temperatures and rates.
[b]Weight of air not considered in computation of energy.
[c]Water-activated. Mass of water not included in calculation of theoretical energy.
[d]Automatically activated; high-rate discharge; 2- to 20-min rate.
[e]Fused salt; heat activated; high-rate discharge; <1- to 60-min discharge times; practical value includes mass of pyrotechnic heat source.
[f]Beta alumina solid electrolyte, sodium polysulfides as fused-salt electrolyte, 350°C operation.
[g]Similar to sodium/sulfur; uses $NaAlCl_4$ as fused-salt electrolyte and $NiCl_2$ as positive reactant. Temperature of operation can be as low as 200°C.
[h]Fused-salt electrolyte, 400°C operation.
[i]Specific energy values based on 0.5 Li per $LiCoO_2$. Li-to-C ratio taken as 1:6.

for representative voltages and specific energies for practical batteries. There are many reasons for the reduction. The weights of nonreactive components (e.g., containers, separators, electrolyte) will degrade the specific energy and performance of any battery system. There may be other, less obvious, factors as well. Consider a "room temperature" system as an example. If, as is so often the case, its performance, because of self-discharge, is degraded significantly by a modest increase in temperature, and if the application is one involving high power, such as use in an electric vehicle, a major weight penalty must be added for the heat-exchange system required to keep temperature below a practical limit. There may well be other temperature-related considerations; an inability to produce required power when below room temperature may require heating prior to use with attendant delays and weight penalties. Other factors influencing the voltage and capacity of a battery are as follows:

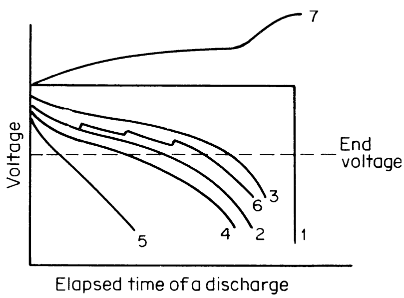


FIGURE 11-26 Battery-discharge characteristics.

Voltage Level. When a battery is discharged, its voltage is lower than the theoretical voltage. The difference is caused by *IR* losses due to cell resistance and by polarization of the active materials during discharge. This is illustrated in Fig. 11-26. The theoretical discharge curve of a battery is shown as curve 1. In this case, the discharge of the battery proceeds at the theoretical voltage until the active materials are consumed and the capacity fully utilized. The voltage then drops to zero. Under actual conditions, a typical discharge curve is similar to curve 2. The initial voltage is lower than theoretical, and it drops off as the discharge progresses.

The Current Drain of the Discharge. As the current drain of the battery is increased, the *IR* loss increases, the discharge is at a lower voltage, and the discharge duration is usually reduced (curve 5). At extremely low current drains, it is possible to approach the theoretical capacities (in the direction of curve 3). In a very long discharge period, the chemical degradation during the discharge becomes a factor and causes a reduction of capacity.

Voltage Regulation. The voltage regulation required by the equipment is most important. As is apparent by the curves in Fig. 11-26, design of equipment to operate to the lowest possible end voltage results in the highest capacity and longest service life. Similarly, the upper voltage limit of the equipment should be established to take full advantage of the battery characteristics. In some applications, where only a narrow voltage range can be tolerated, voltage regulators may have to be employed to take full advantage of the battery's capacity. If a secondary battery is used in conjunction with another energy source, which is permanently connected in the operating circuit, allowances must be made for the voltage required to charge the battery, as illustrated in curve 7, Fig. 11-26. The maximum voltage available from the charger must exceed the maximum battery charge voltage.

11.8.2 Primary Batteries

General. A number of different types of primary (nonrechargeable) batteries are used widely in civilian, industrial, and military applications. They are a convenient, usually relatively inexpensive, lightweight source of power for portable electric devices. The general advantages of primary batteries are reasonably good shelf life, high energy densities at low to moderate rates, little, if any, maintenance, and ease of use. Typical characteristics and applications of these batteries are shown in Tables 11-11 and 11-12.

Leclanche Cell ($Zn\text{-}MnO_2$). The Leclanche or carbon-zinc dry cell, known for over 100 years, is still the most widely used of all the dry-cell batteries because of its low cost, reliable performance, and ready availability. Cells and batteries of many sizes and characteristics have been manufactured to meet the requirements of a wide variety of applications. Characteristics of typical cells are given in Table 11-12.

Composition. The Leclanche cell uses a zinc anode, a manganese dioxide cathode, and an electrolyte of ammonium chloride and zinc chloride dissolved in water. Powdered carbon (acetylene

TABLE 11-11 Major Characteristics and Applications of Primary Batteries

System	Characteristics	Applications
Zinc-carbon (Leclanche) ($Zn\text{-}MnO_2$)	Popular common low-cost primary battery, available in variety of sizes	Flashlight, portable radios and electronics, toys, novelties, instruments, etc.
Magnesium ($Mg\text{-}MnO_2$)	High-capacity primary battery, long shelf life	Military receiver-transmitters, aircraft emergency transmitters
Mercury (Zn-HgO)	Highest capacity (by volume) of conventional types, flat discharge, good shelf life	Hearing aids, medical (heart pacers), photography, detectors, receiver-transmitters, military sensor and detection equipment
Alkaline (Zn-alkaline electrolyte-MnO_2)	Good low-temperature and high-rate performance, moderate cost	Cassettes and tape recorders, calculators, radio, and TV—popular for high-drain primary-battery application
Silver-zinc (Zn-AgO)	Highest capacity (by weight) of conventional types, flat discharge, good shelf life	Hearing aids, photography, electric watches, missiles and space application (larger sizes)
Lithium (lithium-SO_2)	New battery system—recent development; highest-performance primary battery, excellent low-temperature performance, long shelf life	Will have wide, general-purpose application when available; first uses will be military and special civilian applications needing high-capacity and low-temperature performance
Solid electrolyte	Extremely long shelf life, low-power battery	Medical electronics, memory circuits, fusing

black) is mixed with the depolarizer to improve conductivity and retain moisture. As the cell is discharged, the zinc is oxidized and the manganese dioxide reduced. The overall cell reaction is

$$Zn + MnO_2 \rightarrow ZnO + Mn_2O_3$$

Construction. The Leclanche cell is made in many shapes and designs, but in two basic constructions: cylindrical and flat. Similar chemical ingredients are used in both constructions. In the common cylindrical cell (Fig. 11-27) a zinc can serves as the cell container and anode. The manganese dioxide is mixed with acetylene black and solid ammonium chloride, wet with a zinc chloride–ammonium chloride electrolyte, and shaped in the form of a bobbin. A carbon rod is inserted into the bobbin. The rod serves as a current collector and is porous enough to permit the escape of gases, which accumulate in the cell, without allowing leakage of electrolyte. The separator is a cereal paste, also wet with electrolyte, which physically separates the two electrodes and provides the means for ion transfer through the electrode. In the newer "paper-lined" cell, an absorbent kraft paper is used as the separator. This provides thinner separator spacing and lower internal resistance.

Another cyclindrical cell is the "inside-out" construction shown in Fig. 11-28. In this cell, an injection-molded, impervious, inert carbon wall serves as the container of the cell and as the current collector. The zinc anode, formed in the shape of vanes to increase its surface area, is located inside the cell and surrounded by the cathode mix. This ensures efficient zinc consumption and, since zinc is not used as a container, a high degree of leakage resistance.

The flat-cell construction is illustrated in Fig. 11-29. In this cell, carbon is coated on a zinc plate to form a duplex electrode—a combination of the zinc of one cell and the carbon of the adjacent one. The flat cell has a higher energy-to-volume ratio, since the rectangular shape utilizes the available volume better than does the cylindrical cell.

Zinc Chloride Cell. A modification of the Leclanche cell is the zinc chloride electrolyte cell. The construction is similar to the conventional carbon-zinc cell, but the electrolyte contains only zinc

TABLE 11-12 Approximate Service Capacity of American National Standards Institute Sizes of Cylindrical and Flat Carbon-Zinc Cells at Various Current Drains

USASI cell capacity, size	Starting drain, mA	Service capacity, h	USASI cell size	Starting drain, mA	Service capacity, h	USASI cell size	Starting drain, mA	Service capacity h
N	1.5	275	G	15	820	F24	1	475
	7.5	52		75	150		5	150
	15	24		150	65		10	72
AAA	2	290	6	50	700	F30	1.3	275
	10	45		250	150		6.5	40
	20	17		500	70		13	16
AA	3	350	F15	0.4	210	F40	1.3	450
	15	40		2	30		6.5	108
	30	15		4	8		13	52
B	5	420	F12	0.5	435	F60	2	190
	25	65		2.5	103		10	40
	50	25		5	51		20	15
C	5	430	F17	0.6	710	F70	3	550
	25	100		3	155		15	150
	50	40		6	75		30	65
D	10	500	F20	0.7	210	F80	3	600
	50	105		3.5	35		15	165
	100	45		7	12		30	72
E	15	400	F22	0.8	475	F90	3	770
	75	70		4	98		15	200
	150	30		8	49		30	90
F	15	520	F25	1	500	F100	5	1000
	75	105		5	105		25	260
	150	45		10	45		50	110

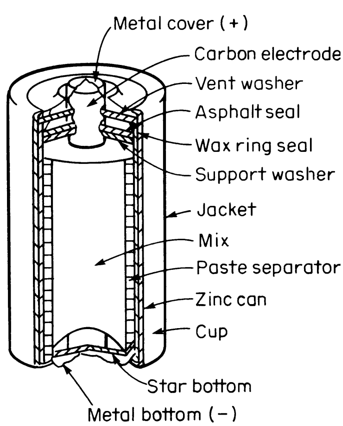


FIGURE 11-27 Cross section of a Leclanche cylindrical cell.

chloride, without the saturated solution of ammonium chloride. The zinc chloride cell is a high-performance cell with improved high-rate and low-temperature performance, and a reduced incidence of leakage. A comparison of the performance of the zinc chloride cell with the conventional cell is presented in Fig. 11-30.

Magnesium Dry Cells (Mg-MnO_2). A magnesium dry cell has two main advantages over the zinc-carbon cell: twice the capacity or service life of an equivalently sized zinc cell and the ability to retain this capacity during storage, even at elevated temperatures (Fig. 11-31). The construction of the magnesium dry cell is similar to the cylindrical carbon-zinc cell except that a magnesium alloy is used instead of zinc. The cathode consists of an extruded mix of manganese dioxide, acetylene black (to provide conductivity and moisture absorbency), magnesium perchlorate electrolyte, barium and lithium chromate as corrosion inhibitors, and magnesium hydroxide as a buffering agent to improve storage life. The degree of "wetness", or amount of water, is critical because water participates in the anode reaction and is consumed during discharge. A carbon rod serves as the cathode current collector. The separator is an absorbent kraft paper as in the paper-lined structure. Sealing of the magnesium cell is critical, since it must be tight to retain cell moisture during storage and also provides a means for the

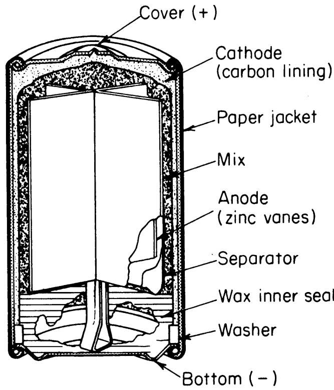


FIGURE 11-28 Cross section of a Leclanche "inside-out" cell.

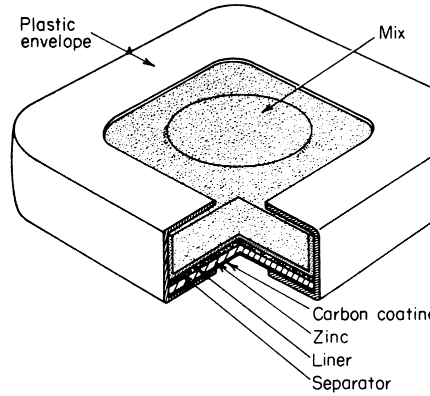


FIGURE 11-29 Leclanche flat cell.

escape of hydrogen gas which forms as the result of a parasitic reaction during the discharge of the battery. This is accomplished by a mechanical vent—a small hole in the plastic top seal washer under a retainer ring, which is deformed under pressure, releasing the excess gas. Magnesium batteries have not been fabricated successfully in flat-cell designs.

The overall reaction of the Mg-MnO_2 cell is

$$Mg + 2MnO_2 + H_2O \rightarrow Mn_2O_3 + Mg(OH)_2$$

At the same time, hydrogen is generated as the result of the parasitic magnesium corrosion reaction:

$$Mg + 2H_2O \rightarrow Mg(OH)_2 + H_2$$

The efficiency of the magnesium anode during a typical discharge is about 70%. Considerable heat is generated during the discharge of a magnesium battery owing to the exothermic side reaction and the *IR* loss resulting from the difference between the theoretical and operating voltages. This heat can be used to advantage at low ambient temperatures to maintain the battery at a warmer temperature.

The good shelf life of the magnesium cell results from a protective film that forms on the inside of the magnesium can, preventing corrosion. This film, however, is responsible for a voltage "delay"—a delay in the cell's ability to deliver full output voltage after it has been placed under load (Fig. 11-32).

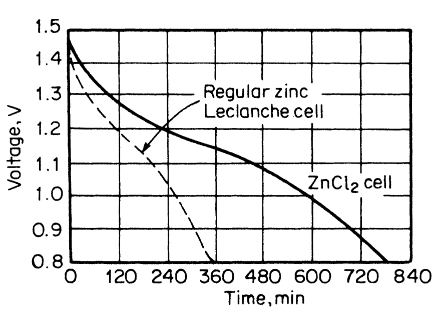


FIGURE 11-30 Comparative performance of zinc chloride and Leclanche cells.

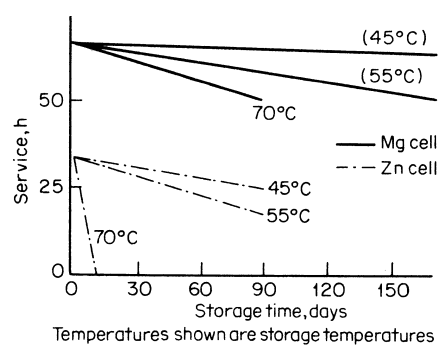


FIGURE 11-31 Comparison of service life vs. storage time of magnesium and Leclanche cells.

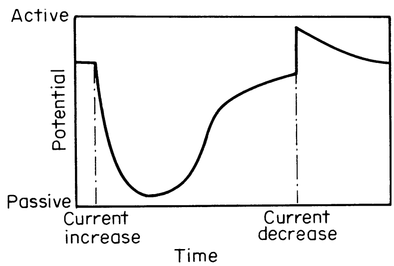


FIGURE 11-32 Voltage delay characteristics of magnesium cells.

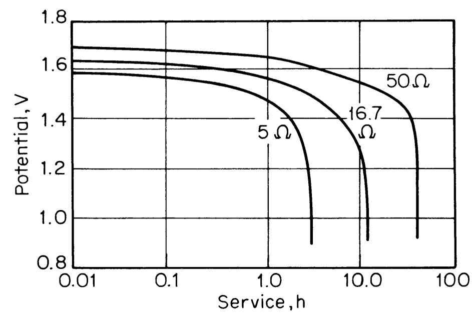


FIGURE 11-33 Discharge curves of magnesium cells (A size, 20°C).

This delay is usually less than 0.3 s, but is longer for discharges at low temperatures and high current drains and after prolonged storage at high temperatures.

Typical discharge curves are given in Fig. 11-33. The magnesium battery is less sensitive to discharge rate than the carbon-zinc cell. The performance of the magnesium cell for various discharge rates and temperatures is shown in Fig. 11-34.

While successful in military use, the magnesium battery has not found wide commercial use. This is due to the fact that the magnesium battery loses its excellent storageability after being partially discharged, and hence is unsatisfactory for long-term intermittent use. Other influencing factors are the higher unit cell voltage and the evolution of hydrogen and heat on discharge which present a potential safety hazard.

Zinc–Mercuric Oxide Cells (Zn-HgO). The zinc alkaline–mercuric oxide battery is noted for its high capacity per unit volume, a relatively constant output voltage during its discharge, and good storage characteristics.

Composition. The zinc–mercuric oxide cell uses amalgamated zinc as the anode, mercuric oxide (mixed with 5% to 10% graphite) as the cathode, and potassium hydroxide as the electrolyte. A saturated solution of zinc oxide is added to the electrolyte to retard the corrosion of zinc, minimize the production of hydrogen, and improve the stability of the cell.

The overall chemical reaction during discharge is

$$Zn + HgO \rightarrow ZnO + Hg$$

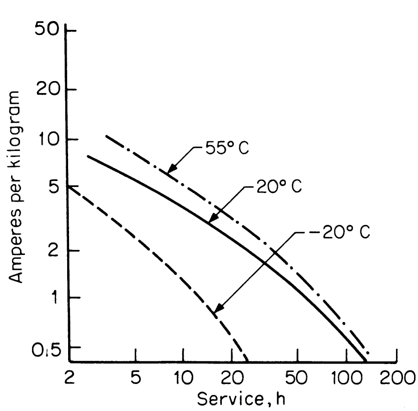
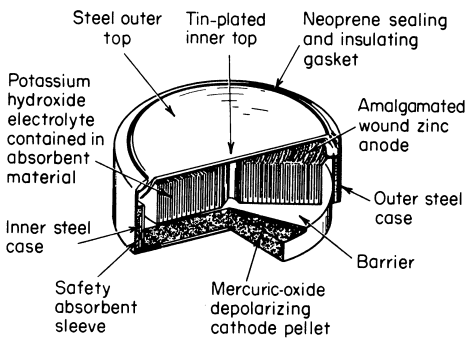


FIGURE 11-34 Service capacity of magnesium cell.

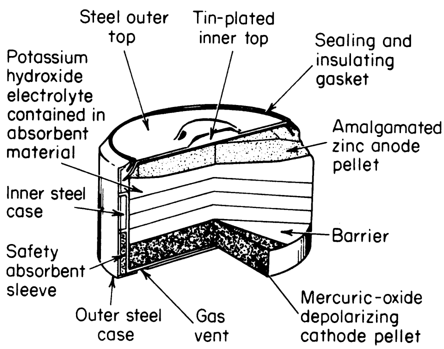
The amperehour capacities of the mercuric oxide cathode and the zinc anode are equalized or balanced, and on completion of the cell's discharge, no residual unoxidized zinc remains. Without this feature, the zinc would continue to react, generating hydrogen gas in the cell.

Construction. The zinc–mercuric oxide cell is manufactured in three basic structures: the wound-anode type, the flat-pressed powdered cathode and anode type, and the cylindrical pressed powdered electrode type.

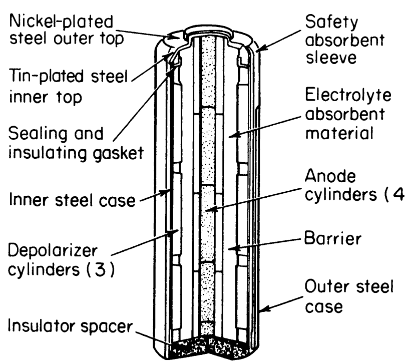
The three types of structures are shown in Fig. 11-35. All constructions use a steel can, which does not take part in the electrochemical reaction and is not consumable, for the cell container. In the wound-anode type, the anode is composed of a corrugated zinc strip with an absorbent paper wound in an offset manner so



Type No.	Max. diam., cm	Height, cm	Weight, g	Rated capacity, mAh
RM 640	1.587	0.965	9.68	360
RM 3	2.498	1.37	22.56	1,500
RM 1438	3.71	1.003	36.22	2,700
RM 1450	3.71	136	51.80	4,500
RM 2550	6.58	1.394	165.20	13,000



Type No.	Max. diam., cm	Height, cm	Weight, g	Rated capacity, mAh
RM 312	0.87	0.34	0.56	36
RM 575	1.143	0.33	1.4	100
RM 675	1.143	0.54	2.24	160
RM 630	1.549	0.58	4.76	350
RM 640	1.574	1.104	7.84	500
RM 4R	3.02	1.66	40.88	3,400



Type No.	Max. diam., cm	Height, cm	Weight, g	Rated capacity, mAh
RM 24	1.0	4.396	14.0	900
RM 601	1.59	2.857	34.16	1,800
RM 3R	2.489	1.651	28.56	2,200
RM 502	1.358	4.90	30.44	2,400
RM 401	1.133	2.844	11.20	800
RM 1R	1.579	1.638	12.04	1,000
RM 12R	1.519	4.959	30.88	3,600
RM 42R	2.922	6.032	148.33	14,000

FIGURE 11-35 Zinc-mercuric oxide cell structures (*left*) and characteristics (*right*).

that it protrudes at one end and the zinc protrudes at the other end. The zinc is amalgamated with 10% mercury and the paper impregnated with the electrolyte, which causes it to swell and produce a positive contact pressure. The cathode is a pellet made of powdered mercuric oxide and graphite which is pressed into the steel can. An absorbent KOH-resistant separator is placed between the two electrodes. The cell has a crimped seal; the can is separated from the top by an insulating neoprene or plastic grommet.

In the pressed-powder cells, the zinc powder is amalgamated and pressed into a pellet with sufficient porosity to allow electrolyte impregnation. A double-can structure is used in the larger-sized cells as a safeguard. Under excessive gas pressures, the compression of the upper part of the grommet by internal pressure allows the gas to escape into the space between the two cases. A paper tube surrounds the inner can so that any liquid carried by the discharging gas will be absorbed, maintaining a leak-resistant structure. Release of the excess gas pressure reseals the cell.

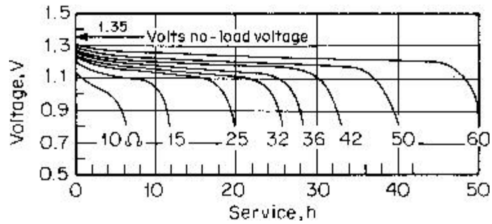


FIGURE 11-36 Discharge curves of zinc–mercuric oxide cells.

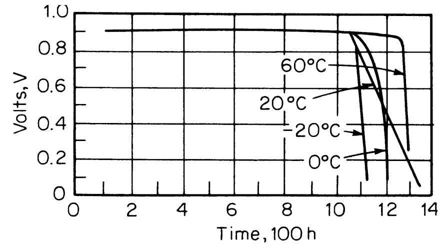


FIGURE 11-37 Discharge curves of cadmium-mercuric oxide cells.

The cylindrical and button cells are variations of the pellet design. The button cell uses a gelled electrolyte to reduce electrolyte leakage further. The mechanical and electrical specifications of representative cells of each of the three constructions are given in Fig. 11-35.

Voltage. The open-circuit voltage of the mercury cell is 1.35 V and is reproducible within 0.001 V. On discharge, the cell is characterized by a very flat discharge as shown in Fig. 11-36. The cutoff voltage usually is 0.9 to 1.0 V per cell.

Cadmium–Mercuric Oxide Cell (Cd-HgO). The substitution of cadmium for zinc results in a very stable system with a predicted shelf life of up to 10 years, as well as performance at low temperatures (Fig. 11-37). The watthour capacity of this cell, because of its lower voltage, is about 60% of zinc–mercuric oxide cell capacity. In design, the cadmium–mercuric oxide cell is similar to the zinc–mercuric oxide cell. These mercuric-oxide systems (Zn and Cd) are being evaluated because of environmental concerns.

Zinc–Silver Oxide Cell (Zn-AgO). The primary zinc alkaline–silver oxide cell is similar in design to the small zinc–mercuric oxide button cell but uses silver oxide in place of mercuric oxide. Cells range in capacity up to 200 mAh for use at 50-h and lighter loads. The silver oxide cell has an open-circuit voltage of 1.6 V and operates about 0.2 V higher than the mercuric oxide cell. Typical discharge curves for this cell are given in Fig. 11-38. The silver oxide cell has a higher energy density (on a weight basis) and is less sensitive to a reduction in ambient temperature than the mercuric oxide cell. At the design loads, the cell will deliver about 70% of its 20°C performance at 0°C and 35% at –20°C. These characteristics make this battery desirable for use in hearing aids, photographic applications, and electronic watches.

Alkaline–MnO_2 Cell (Zn-MnO_2). The zinc alkaline–MnO_2 cell uses the same electrochemically active materials, zinc and manganese dioxide, as the Leclanche cell, but differs in construction and in the use of highly conductive potassium hydroxide electrolyte which results in a lower internal resistance. The advantage on low-rate or intermittent discharge is marginal, but on high- and continuous-drain conditions, the alkaline cell can deliver from 2 to 10 times the amperehour capacity of the Leclanche cell. Its performance at low temperatures is superior to other commercially available dry batteries.

$$Zn + 2MnO_2 \rightarrow Mn_2O_3 + ZnO$$

The electrolyte undergoes no change during the discharge, maintaining its high conductivity throughout the cell's life. It thus differs from the Leclanche cell, whose resistance increases during the discharge.

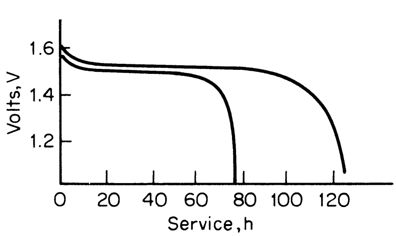


FIGURE 11-38 Discharge curves of zinc–silver oxide cells.

Construction. The principal features of the Zn-MnO_2 cell are the manganese dioxide cathode of high density, a zinc anode of high surface area, and the highly conductive potassium hydroxide electrolyte. As illustrated in Fig. 11-39, the MnO_2, mixed with graphite or carbon black, is pressed against the inner surface of the can, which serves as the cathode current collector. The anode is centrally located and consists of a mixture of granular or powdered zinc, which is amalgamated to reduce hydrogen evolution, and the electrolyte. In some designs, a gelling agent is used to immobilize the electrolyte and minimize leakage. A highly absorbent, chemically inert material separates the electrodes.

Voltage. The open-circuit voltage of the alkaline-MnO_2 cell is 1.5 V. Its discharge is similar to the Leclanche cell, but it is superior at the heavier discharge loads. Typical discharge curves are given in Fig. 11-40.

Service Life. At light loads, the service life of the alkaline cell is about the same as the Leclanche. However, its service capacity remains relatively constant with increasing load and is much superior to the Leclanche at higher current drains. The performance of the alkaline cell for various discharge rates and temperatures is shown in Fig. 11-41.

Effect of Temperature. The alkaline-MnO_2 cell performs well at low temperatures, excelling over the best Leclanche cells. The cell operates to temperatures as low as –40°C.

Shelf Life. The shelf life of the alkaline-MnO_2 cell is moderately superior to the Leclanche cell. Capacity retention is about 90% after 1 year of storage at 20°C.

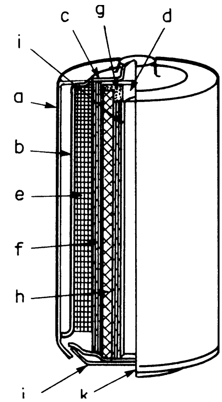


FIGURE 11-39 Construction of alkaline-MnO_2 cells: (*a*) outer nickel-plated can; (*b*) tube adapter; (*c*) inner gold-plated can; (*d*) insulator disk; (*e*) depolarizer; (*f*) outer absorbent with barrier; (*g*) insulating ring; (*h*) anodes; (*i*) inner absorbent; (*j*) molded double top; (*k*) clear plastic dielectric jacket. Electrolyte not shown. (*Duracell.*)

Lithium Primary Batteries. The lithium battery has a number of advantages over other primary-battery systems. Lithium is an attractive anode because of its reactivity, light weight, and high voltage; cell voltages range between 2 and 3.6 V, depending on the cathode material.

The advantages of the lithium battery include high specific energy and energy density, on the order of 250 Wh/kg and 400 Wh/dm^3, respectively; high power density; flat discharge characteristics; excellent service over a wide temperature range, down to –40°C or below; and good shelf life (up to 5 years without refrigeration is anticipated).

Nonaqueous solvents must be used as the electrolyte because of the reactivity of lithium with aqueous solutions. Organic solvents, such as acetonitrile and propylene carbonate, and inorganic solvents, such as thionyl chloride, are typical. A compatible solute is added to provide the necessary

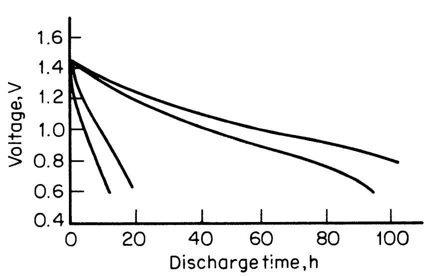


FIGURE 11-40 Discharge curves of an alkaline-MnO_2 cell at 20° C.

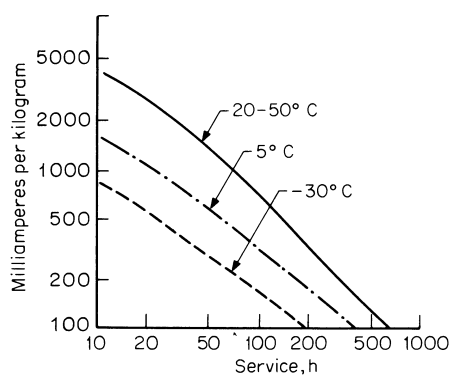


FIGURE 11-41 Service capacity of an alkaline-MnO_2 cell.

electrolyte conductivity. A number of different materials—sulfur dioxide, carbon monofluoride, vanadium pentoxide, and copper sulfide—are used as the active cathodes.

Lithium–Sulfur Dioxide Cell. One advanced lithium primary cell uses sulfur dioxide for the cathode material and an electrolyte consisting of acetonitrile and lithium bromide. The cell reaction is

$$2Li + 2SO_2 \rightarrow Li_2S_2O_4 \text{ (lithium dithionite)}$$

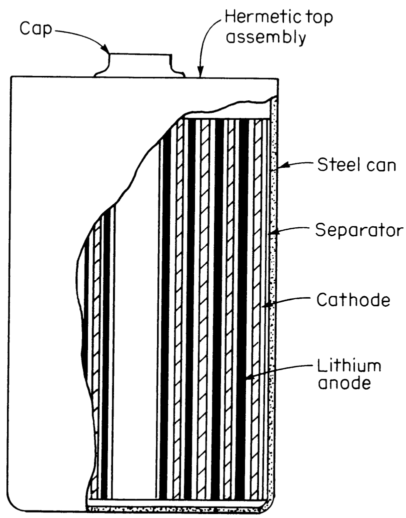


FIGURE 11-42 Cross section of a cylindrically wound Li-SO_2 cell.

The cell is typically fabricated in a cylindrical structure as shown in Fig. 11-42. A "jellyroll" construction is used, made by spirally winding strips of lithium ribbon, a polypropylene separator, and the cathode electrode (a Teflon-carbon mix pressed on an aluminum screen). This design provides the high surface area and low cell resistance necessary to obtain high-current and low-temperature performance. The roll is inserted in a steel container which is electrically connected to the anode. The cathode, in turn, is connected to its terminal and the cell hermetically sealed. The electrolyte-depolarizer mixture (70% SO_2) is added through a temporary opening in the cell to an SO_2 pressure of about 4 atm. The critical manufacturing operations are carried out in dry rooms or dry boxes to minimize the moisture content of the cell.

The good shelf life of the lithium-SO_2 cell is attributed to the protective film formed by the initial reaction of lithium and SO_2, which prevents further reaction or loss of capacity during stand. During discharge, the SO_2 is depleted and the cell pressure reduced. The discharge is generally terminated by the deactivation of the carbon electrode by blocking the active area from precipitation of the discharge product.

The performance characteristics of the Li-SO_2 cell are given in Figs. 11-43 through 11-46. Figure 11-43 shows typical discharge curves for the cell at various discharge loads. The high cell voltages and flat discharge curves shown are characteristic of this cell.

Figure 11-44 presents the performance of the Li-SO_2 cell at various temperatures and discharge rates. The superior low-temperature performance (over 60% of the normal-temperature performance available at –30°C) is noteworthy. Figure 11-45 illustrates the shelf life of the lithium cell. Although

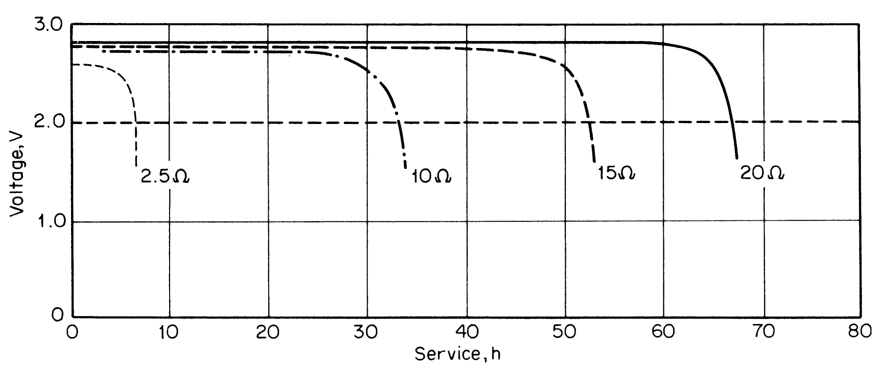


FIGURE 11-43 Typical discharge curves of Li-SO_2 cell at 20°C at various loads.

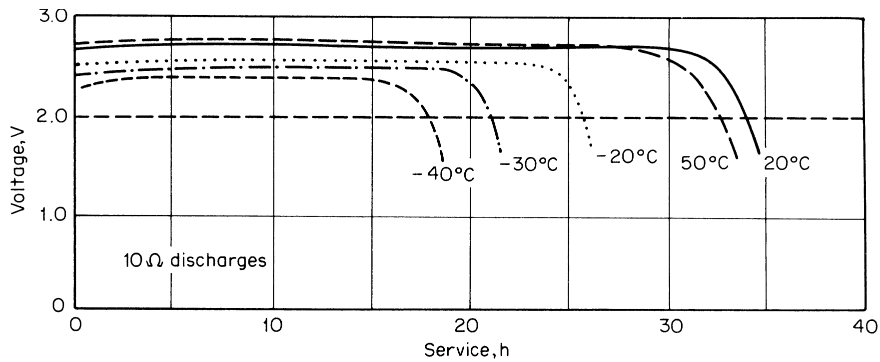


FIGURE 11-44 Effect of temperature on performance of Li-SO_2 cells (D size).

very long-term storage has not been demonstrated experimentally, the data suggest a long-shelf-life capability even at high temperatures, particularly with newer cell designs.

Special attention must be given to the design and use of the lithium battery, since it contains materials that are potentially flammable and toxic. Properly designed cells are equipped with safety vents which release SO_2 when the cells reach high temperatures and pressures, thus preventing explosive damage.

The lithium battery can deliver unusually high current. Since high internal temperatures can develop from continuous high current drain, short circuiting, or inadvertent internal cell shorts, such use must be avoided. It is advisable to equip batteries with fuses to protect against short circuiting.

Charging lithium-SO_2 cells may result in explosion of cells (even those which are equipped with safety vents) and should not be attempted. Similarly, cells or groups of cells should not be connected in parallel without diode protection to prevent one group of

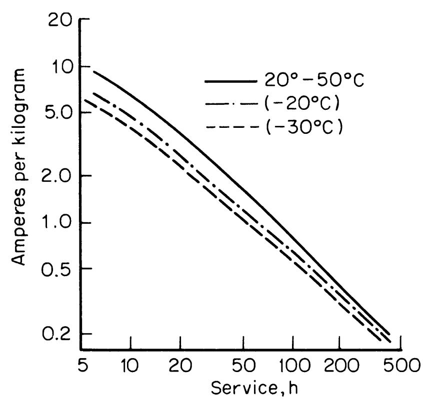


FIGURE 11-45 Service capacity of Li-SO_2 cells at various temperatures.

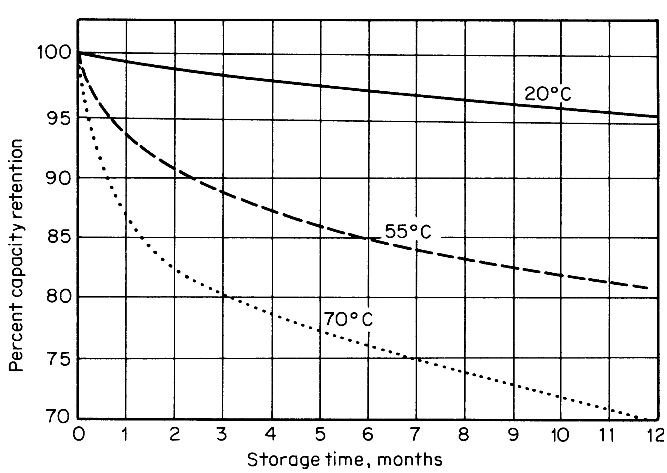


FIGURE 11-46 Shelf life of Li-SO_2 cell at different storage temperatures.

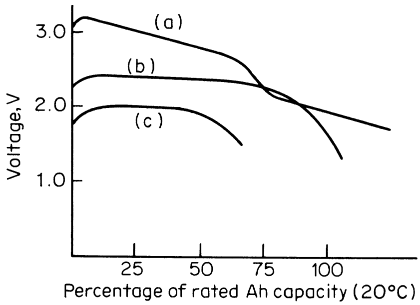


FIGURE 11-47 Typical discharge curves of lithium–solid cathode cells: (*a*) vanadium pentoxide, 8-h rate, 20° C; (*b*) carbon monofluoride, 20-h rate, 20°C; (*c*) carbon monofluoride, 20-h rate –20°C.

cells from charging the other. Forced discharging, which could occur with cells that are connected in series or to an external power supply, also may result in venting and/or explosion. Currently, special procedures govern the transportion, shipment, and disposal of lithium batteries. While it requires special handling and design, the many advantages of the lithium cell result in increasing use of this battery in military and civilian areas.

Other Lithium–Organic Electrolyte Cells. A number of other cathode materials, mostly solid, have been developed for lithium primary cells. Cells using these solid cathodes have the advantage of being nonpressurized but do not have the high current capability of the SO_2 system. Some cells are specifically designed for low-rate applications, using "bobbin"-type constructions. These cells deliver higher energy outputs and should provide safe operation, since they are self-limiting in energy and current output. Typical discharge curves for the carbon monofluoride and vanadium pentoxide cells are shown in Fig. 11-47.

Inorganic Electrolyte Cells. Certain nonaqueous inorganic liquids, such as thionyl chloride ($SOCl_2$) and sulfuryl chloride (SO_2Cl_2), also are capable of forming a passivating film on the lithium anode, similar to that in SO_2, and can be used both as the electrolyte solvent and as the active cathode material. A commonly used solute is $LiAlCl_4$. Cells made with these components are similar in construction to the Li-SO_2 cell but are not pressurized at room temperature because of the lower vapor pressure of the thionyl chloride. The Li-$SOCl_2$ cells also operate at approximately 0.5 V higher than the comparable Li-SO_2 cell. Figure 11-48 compares the discharge curves of the two batteries and illustrates the higher voltage and energy output of the thionyl chloride cell. The cells exhibit good shelf life, but the passivating protective film that forms, particularly at high-temperature storage, is not readily penetrated, and excessively long voltage delays occur, especially at high-rate and low-temperature discharges. Lithium–thionyl chloride cells are available in a range of capacities from 1.4 to 8,000 Ah. Corresponding specific energy ranges from 240 to 450 Wh/kg at

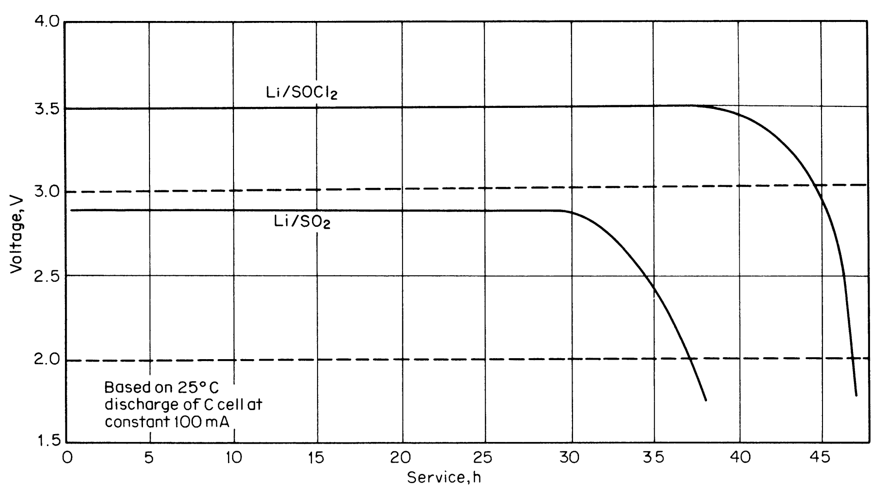


FIGURE 11-48 Discharge curves of Li-$SOCl_2$ and Li-SO_2 cells.

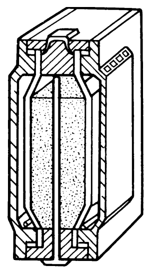


FIGURE 11-49 Construction of primary zinc-air cell. (*Duracell.*)

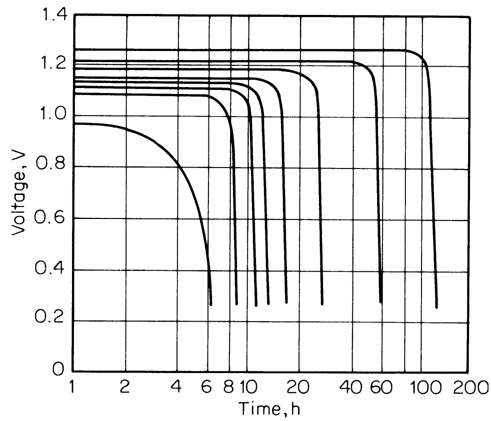


FIGURE 11-50 Typical discharge curves of zinc-air primary cells. (*Duracell.*)

rates of about 10 days. The cells display safety characteristics that reflect engineering to accommodate the high-energy nature of the reactants. A representative cell reaction is

$$2Li + SOCl_2 \rightarrow 2LiCl + {}^1/_2S + {}^1/_2SO_2$$

Zinc-Air Cells. Primary batteries using air as the depolarizer can deliver high energy densities, since they do not contain active cathode material. Zinc-air batteries, using bulky zinc anodes, a carbon-air cathode, and potassium hydroxide electrolyte, constructed in heavy-glass or hard-rubber containers, had been used successfully for many years in railway signals and similar applications. They were noted for their high energy densities but low power output capability. Lower-capacity zinc-air batteries, up to the 20-Ah size, are now being developed, using thin Teflon-coated fuel cell-type electrodes. These new structures have high specific energy, on the order of 220 Wh/kg, and are capable of moderately high current drains.

Construction. A typical construction of the primary zinc-air cell is shown in Fig. 11-49. It consists of a high-surface-area, gelled zinc anode, fabricated by pressing zinc granules with a gelled potassium hydroxide electrolyte, and a teflonated air cathode made of a low-cost, non-noble-metal catalyst. Cylindrical-shaped and button cells also have been designed. Effective sealing is essential in this construction to prevent electrolyte leakage. The cells are stacked to form a battery; space must be left between each cell for air circulation.

Performance Characteristics. The zinc-air battery is best suited for continuous, moderate-drain discharges at temperatures between –20 and +35°C. Intermittent operation usually results in a loss of capacity due to the drying out of the cell. Since the cell depends on oxygen from the air for its operation, differences in performance occur with variation in air circulation. Figure 11-50 shows typical curves for continuous discharge at 20°C. The flat discharge is characteristic of this cell. Figure 11-51 shows the energy output of the zinc-air battery at various temperatures. In addition to the reduction of service life, the average voltage decreases with decreasing discharge temperature.

Primary zinc-air cells must be stored in sealed bags after manufacture to maximize storage life. Limits of storage as well as the integrity and leak resistance of the cell seal

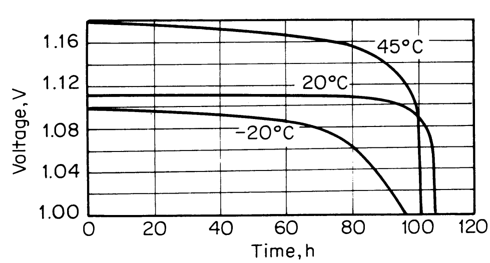


FIGURE 11-51 Voltage vs. time curve of zinc-air cell at various temperatures. (*Duracell.*)

TABLE 11-13 Characteristics of Solid Electrolyte Cells

System	Cell voltage, V	At 100-h rate Wh/dm^3	At 100-h rate Wh/kg
$Ag/RbAg_4I_5/I_2$	0.66	40–80	15–25
$Li/LiI(Al_2O_3)/PbI_2$, Pb	1.9	100–200	35–70
$Li/LiI(Al_2O_3)/PbI_2$,PbS,Pb	1.9	300–500	75–150
$Li/LiI/I_2$ (poly-2-vinylpyridine)	2.8	250–500	120–180

have yet to be determined. Once the packaging has been removed, the battery should be put into service shortly thereafter, since moisture loss results in dry-out and reduced capacity.

Solid Electrolyte Batteries. Most batteries depend on the ionic conductivity of liquid electrolytes for their operation. The solid electrolyte batteries depend on the ionic conductivity of an electronically nonconductive salt in the solid state as, for example, Ag^+ ion mobility in silver iodide. Cells using these solid electrolytes are low-power (microwatt) devices but have an extremely long shelf life and the capability of operating over a wide temperature range. The absence of liquid eliminates corrosion and gassing and permits the use of a hermetically sealed cell. The solid electrolyte batteries are used in medical electronics (in devices such as heart pacemakers), for memory circuits, and for fusing and other such applications requiring a long-life, low-power battery.

Several types of solid electrolyte batteries are being marketed using different solid electrolytes and active materials. The characteristics of several of the available types are summarized in Table 11-13. Of special significance are the high energy densities (5 to 10 Wh/in^3) achieved with the Li-anode solid electrolyte battery. Typical construction of solid electrolyte cells is shown in Fig. 11-52. The design features a sealed structure to exclude moisture and maintain a high-density, void-free package.

The discharge curves for various solid electrolyte cells at 25°C are given in Fig. 11-53. Since the batteries are designed primarily for low current drain, continuous discharge at high rates is not practical. The energy density and power density also are dependent on the operating temperature. A significant characteristic of the solid electrolyte battery is its long shelf life. Projections based on limited tests (1 to 2 years) predict a shelf life exceeding 15 years at 20°C. Figure 11-54 shows the projected capacity retention at various storage temperatures.

Other Primary Batteries. Many other electrochemical systems have been used as primary batteries to obtain special performance characteristics. The more prominent ones are listed in Table 11-14.

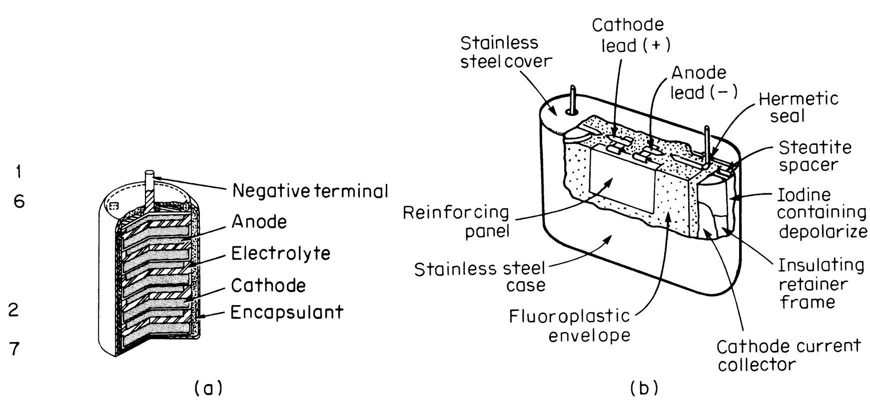


FIGURE 11-52 Construction of solid electrolyte cells: (*a*) silver iodide cell; (*b*) lithium iodide cell.

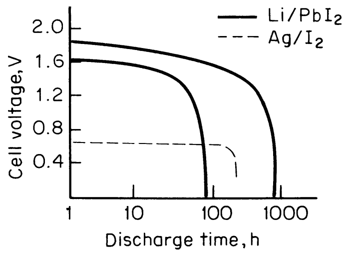


FIGURE 11-53 Typical discharge curves of solid dielectric electrolyte cells.

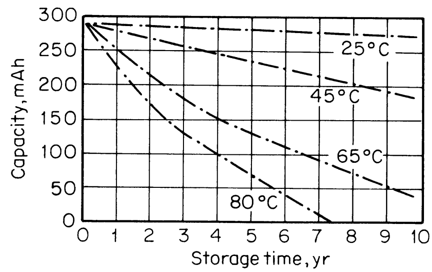


FIGURE 11-54 Capacity retention vs. temperature of solid electrolyte cells. (*Sigma Technologies International, Inc.*)

Recharging Primary Batteries. Recharging primary batteries is a practice that should generally be avoided since the cells are not designed for such use. In most instances it is impractical, and it could be hazardous in cells that are tightly sealed and not vented to permit the release of gases that form during charge and discharge. Most primary cells and batteries are labeled with a cautionary notice advising that they should not be recharged.

TABLE 11-14 Major Characteristics and Applications of Secondary Batteries

System	Characteristics	Applications
Lead-acid		
Automotive	Popular, low-cost secondary battery—moderate capacity, high-rate and low-temperature performance	Automobile starting, lighting, ignition (SLI); lawnmowers, tractors, marine, float service
Motive power	Designed for deep 6- to 9-h discharge, cycling service	Industrial trucks, materials handling; special types used for submarine power
Stationary	Designed for standby float service, long stand life	Emergency power—uninterruptible power supplies
Valve-regulated	Low maintenance, moderate cost, good float capability moderate cycle life	Many of the above applications, plu, TV, portable tools, lights and appliances, radios and cassettes and tape players
Nickel-cadmium		
Vented	Good high-rate, low-temperature capability; flat voltage, excellent cycle life	Aircraft batteries, industrial and emergency-power applications, communication equipment
Valve-regulated	Good high-rate, low-temperature performance, excellent cycle life, low maintenance	Photography, portable tools, appliances, standby power
Lithium-ion	Good gravimetric and volumetric energy and power, high cell voltage	Consumer electronics
Nickel-metal hydride	Replacement for nickel-cadmium, slightly higher cost, less robust low-temperature peformance	Same as nickel-cadmium
Zinc–silver oxide	High energy density, good high-rate capability, low cycle life	Lightweight portable radio, TV, and communication equipment; torpedo propulsion, drones, submarines, and other military applications

Some Leclanche zinc-carbon cells can be recharged for several cycles under carefully controlled conditions. For successful recharging, the cell should be placed on charge soon after removal from discharge and at a low rate (about 16 h charge time). The cell voltage on discharge should not be less than 1.0 V when it is removed for charging. The cells must be returned to service soon after recharging, since the shelf life after recharge is poor. Note the recent introduction of Rayovac Renewal "reusable alkaline batteries."

11.8.3 Secondary Batteries

General. Secondary (rechargeable) batteries are widely used in many applications. Most are characterized, in addition to their ability to be recharged, by high power densities, by their capability to be discharged at high rates, by flat discharge curves, and by good low-temperature performance. Their energy densities are usually lower than those of primary batteries. The characteristics of some secondary batteries are listed in Table 11-14.

The applications of secondary batteries fall into two major categories:

1. Those applications where the secondary battery is used essentially as a primary battery, but recharged after use. Secondary batteries are used in this manner for convenience (as in handheld calculators or electronic flash units), for cost savings (as they can be recharged rather than replaced), or for power drains beyond the level of primary batteries.
2. Those applications where the secondary battery is used as an energy-storage device, being charged by a prime energy source and delivering its energy to the load on demand. Examples are automotive and aircraft systems, and emergency and standby power sources.

A summary of some of the major applications of the various types of secondary batteries is given in Table 11-14.

Lead–Acid Battery (Pb-PbO$_2$). The lead-acid battery is the most widely used secondary battery. Its low cost, reliability, and generally favorable performance characteristics account for its acceptance in many different applications. This type of battery is manufactured in many sizes, ranging in capacity from less than 1 Ah (small plastic-encased or sealed portable cells) to several thousand amperehours for stationary and vehicle-propulsion types. Characteristics of typical cells are summarized in Table 11-14.

Composition. The lead-acid battery uses highly reactive sponge lead for the negative electrode, lead dioxide as the active positive material, and a sulfuric acid solution for the electrolyte. As the cell discharges, the active materials of both electrodes are converted into lead sulfate. The sulfuric acid electrolyte also takes part in the reaction producing water. On charge, the reverse reactions take place. The state of charge of the battery can be determined in some cases by measuring the specific gravity of the electrolyte, which decreases on discharge and increases on charge. The discharge and charge reactions of the battery are

$$Pb + PbO_2 + 2H_2SO_4 \underset{\text{charge}}{\overset{\text{discharge}}{\rightleftharpoons}} 2PbSO_4 + 2H_2O$$

At the end of charge, electrolysis of water may occur, producing hydrogen at the negative and oxygen at the positive electrode.

Construction. The most common construction for the lead-acid cell is the pasted-plate design. A cross section of an automotive-type battery using this construction is shown in Fig. 11-55. The active material for each electrode is prepared as a paste by mixing finely divided lead oxides and suitable expander materials with sulfuric acid. The paste is spread onto a lead-alloy grid which provides the necessary electrical conductivity and structure to hold the active materials. The resulting plates are soldered or welded to connecting straps to form positive and negative plate groups which are interleaved. Separators are placed between the electrodes, and the completed element is

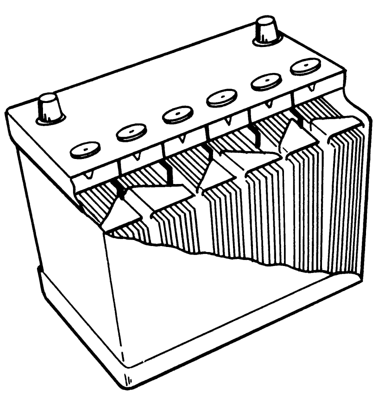


FIGURE 11-55 Cross section of lead-zinc automotive battery. (*GNB, Inc.*)

placed in a container. The container is designed with a sediment space under the element to safely collect any of the active material that is dislodged. Sufficient headroom is provided above the plates to hold excess electrolyte and allow electrode growth. Conventional lead-acid batteries often employ alloy-strengthened grids; calcium and antimony are the common alloying metals. The alloying is necessary to provide adequate strength for the thin grid structure to facilitate casting.

The low maintenance value regulated lead-acid batteries, which became popular in the mid-1970s, use calcium-lead grids, which are more resistant to corrosion and self-discharge. Water loss from gassing during charge is minimized in this design. The battery is filled with an excess of electrolyte and closed to prevent contamination. Other design features are improved separator materials which reduce the possibility of internal shorting, enclosed internal connectors, and lightweight, high-impact-strength polypropylene cases.

Other lead-acid batteries are similar in design to the automotive battery but vary in lead-alloy composition, plate thickness, separators, container, etc., to optimize the performance characteristics for the particular application.

General Performance Characteristics. The general performance characteristics of the lead-acid battery are given in Fig. 11-56.

Voltage. The nominal voltage of the lead-acid cell is 2 V; the voltage on open circuit is a direct function of the specific gravity, ranging from 2.12 V for a cell with 1.28 specific gravity to 2.05 V at 1.21. Figure 11-57 presents typical discharge curves for the lead-acid cell. The end of discharge cutoff voltage is usually about 1.75 V per cell but can be as low as 1.0 V per cell at extremely high rates, as in automotive starting service.

Specific Gravity. The selection of the specific gravity of the electrolyte at the start of discharge depends on the service requirements. The electrolyte concentration must be high enough for good electrical conductivity and to fulfill electrochemical requirements, but not so high as to cause separator deterioration or corrosion of other parts of the cell, which would shorten life and increase self-discharge. A specific gravity of 1.26 to 1.28 is usually used in automotive and high-performance batteries, and one as low as 1.21 for stationary standby batteries. The specific gravity should be reduced in high-temperature climates.

During discharge (Fig. 11-57), the specific gravity decreases about 0.125 to 0.150 points from a fully charged to a fully discharged condition. The change is proportional to the amperehours discharged. The specific gravity is thus an excellent means for checking the state of charge in a flooded battery. One must wait 5 or 10 min after discharging prior to measurement of specific gravity to allow the acid concentration to equilibrate. On charge, the change in specific gravity should similarly be proportional to the amperehour charge accepted by the cell, but there is a lag because complete mixing of the electrolyte does not occur until gassing commences near the end of the charge.

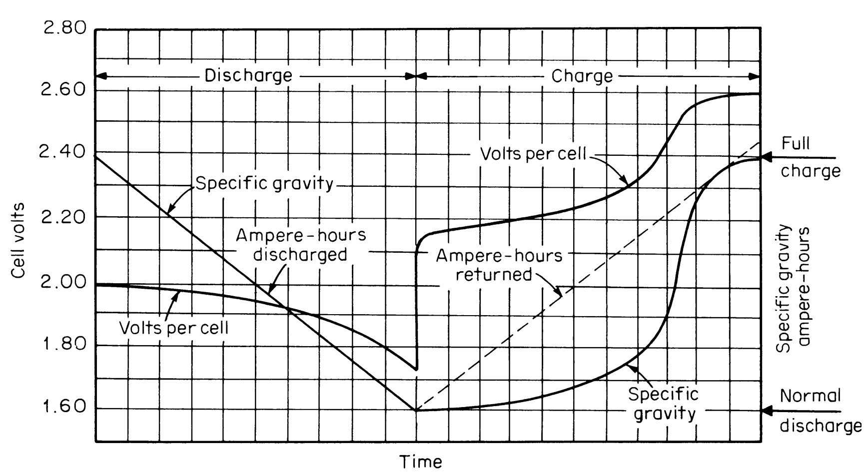


FIGURE 11-56 Performance characteristics of lead-acid cells. (*GNB, Inc.*)

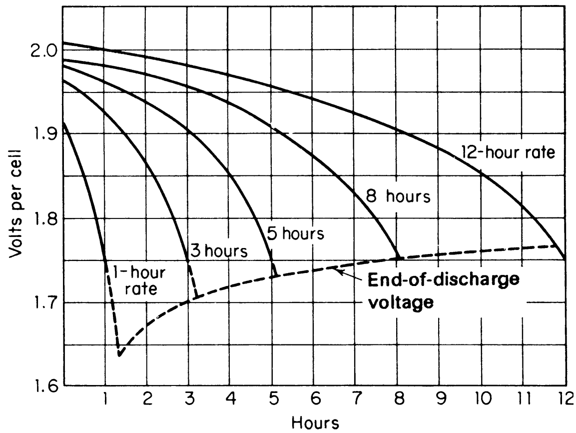


FIGURE 11-57 Discharge curves of lead-acid cells at different hour rates.

Service Life. The service life of a typical automotive-type lead-acid cell is shown in Fig. 11-58 for different discharge rates and temperatures. These curves have been normalized to the 20-h rate at 25°C. A 100-Ah cell, for example, will deliver 20 h of service when discharged at 5 A at 25°C or 1 h of service when discharged at –40°C at 20 A. Typically, higher service capacity is obtained at lower discharge rates and higher temperatures. In general, a battery may be discharged without harm at any rate of current it will deliver, but the discharge should not be continued beyond the point where the cell approaches exhaustion or where the voltage falls below a useful value.

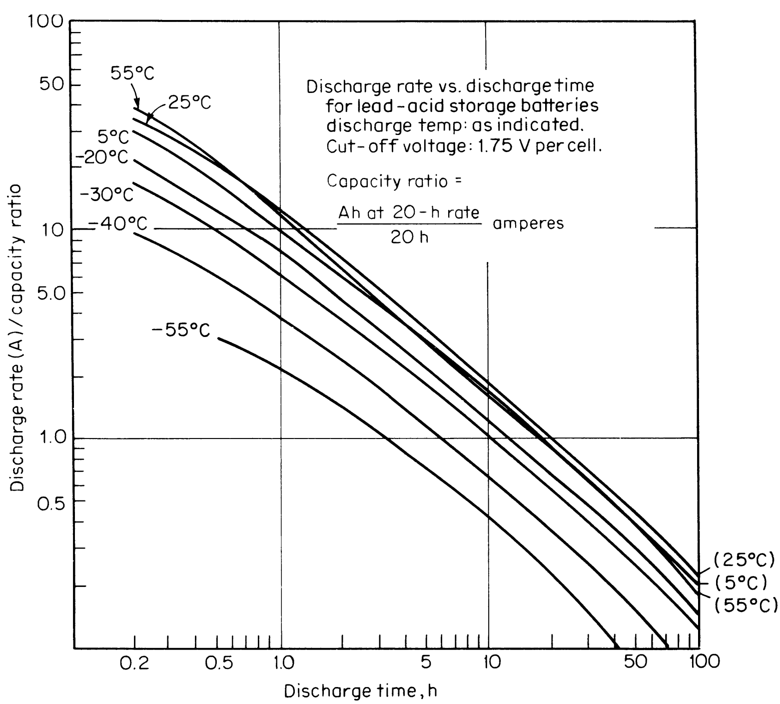


FIGURE 11-58 Service-capacity curves of lead-acid batteries.

Automotive cells are typically made with thinner plates, have a larger surface area, and use a higher concentration of electrolyte than motive-power and stationary batteries, so higher currents can be drawn at higher voltage levels. Hence, the electrical output of stationary batteries, which are designed for long-life service, will be somewhat lower to that indicated in Fig. 11-58.

Temperature. The variation of the performance of the lead-acid cell at different temperatures and loads is given in another form in Fig. 11-59. Although the battery will operate over a wide temperature range, continuous operation at high temperatures may reduce cycle life as a result of the increase in the rate of corrosion.

Self-Discharge. The self-discharge rate (loss of battery capacity during storage) depends on a number of factors, including the type of lead alloy used, the concentration of electrolyte, the age of the battery, and temperature. Self-discharge is caused by local chemical reactions between components of the plate and occurs almost entirely in the negative electrode. The rate of self-discharge is about 15% per month for antimonial-lead batteries at 25°C. Batteries using purer lead grids have substantially lower rates of self-discharge. Capacity lost by self-discharge can be recovered by recharging the battery. For best practice, a battery on stand should be recharged every 3 to 6 months, since prolonged storage and self-discharge can cause irreversible damage and make recharge difficult, owing to sulfation of the plates.

Lead-Calcium Cells. The lead-calcium cell uses a small percentage of calcium to give the lead grid the necessary physical rigidity in place of a larger amount of antimony normally used. Thus, it is closer to a pure-lead grid and has considerably reduced self-discharge due to local chemical action. The calcium alloy cells are best suited for standby float or open-circuit service rather than a cycling type of use. Frequent recharging in cycle use causes the grid in the positive plate to enlarge owing to corrosion of the lead grid and shortens life expectancy. Because of the improved performance, in terms of self-discharge and life, the use of calcium has largely replaced the use of antimony in the storage-battery market.

Nickel-Cadmium Batteries. The major alkaline secondary battery is the nickel-cadmium battery, which is noted for high power capability, long cycle life, good low-temperature performance, ruggedness, and reliability. This battery system is manufactured in many sizes, ranging from the small sealed button and

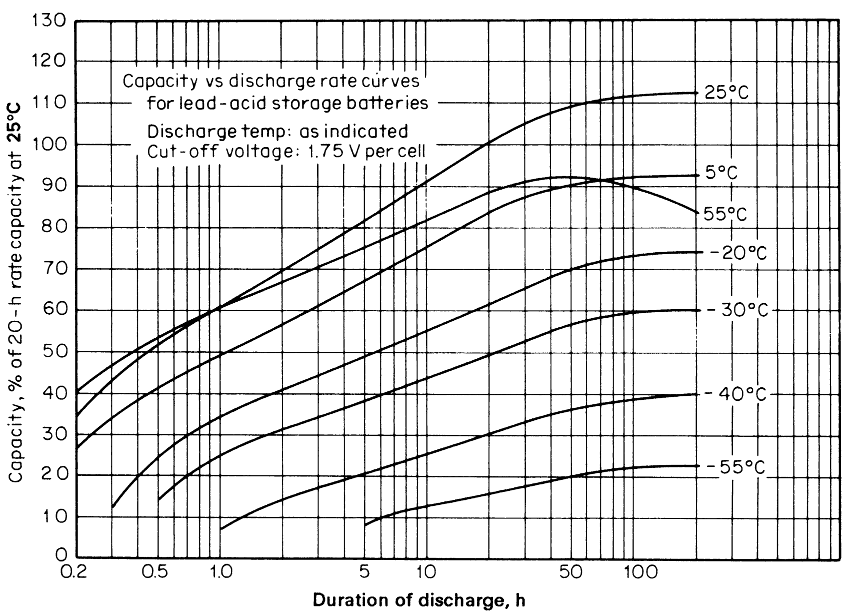


FIGURE 11-59 Performance of lead-acid batteries at various temperatures.

cylindrical cells with capacities as low as 0.020 Ah to larger vented cells for stand by and emergency service with capacities over 1,000 Ah. The nominal voltage of the nickel-cadmium cells is 1.2 V; the open-circuit voltage is 1.4 V. Table 11-14 lists the characteristics of different types of nickel cadmium cells.

Nickel-Metal Hydride. The nickel-metal hydride battery technology has become widely used in the last 10 years for consumer electronics as well as in propulsion and industrial applications. This technology is replacing nickel-cadmium batteries in many applications due to the environmental concerns related to cadmium. Nickel-metal hydride cells use the same nickel electrode as does the nickel-cadmium battery, but the negative electrode consists of hydrogen (generated during charge) absorbed in a metal alloy. This electrochemical couple is well-known in a related technology, nickel-hydrogen batteries, which are used in high value satellite applications. The cost of the metal hydride battery is more than for nickel-cadmium, but still competitive for consumer and industrial products. It also exhibits more capacity per unit mass than does nickel-cadmium, but generally has lower discharge rate capability.

Construction of a typical cylindrical nickel-metal hydride cell for consumer products is shown in Fig. 11-60. The cell contains a safety vent, but is normally sealed to retain the hydrogen that is generated during charge. Cells are also available in button and prismatic configurations. Large, industrial scale cells are also available, and an example is shown in Fig. 11-61.

Typical voltage performance curves for a cylindrical cell are shown in Fig. 11-62. The technology has discharge limitations below about –20°C, but performs reasonably above that temperature. It does exhibit the

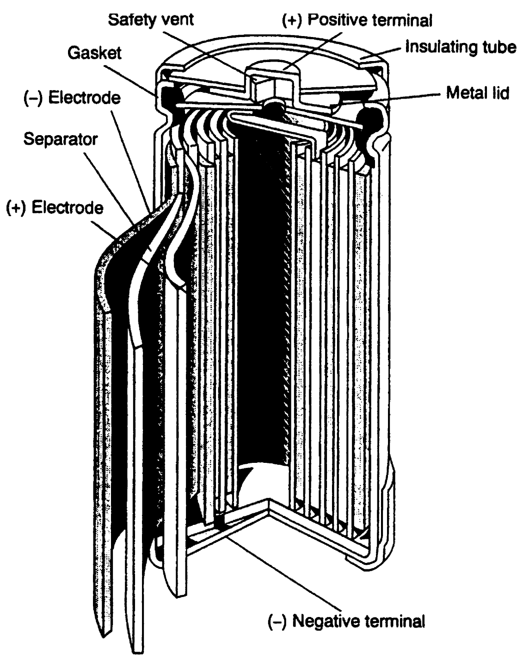


FIGURE 11-60 Construction of a sealed cylindrical nickel-metal hydride battery. (*Courtesy of Duracell Inc.*)

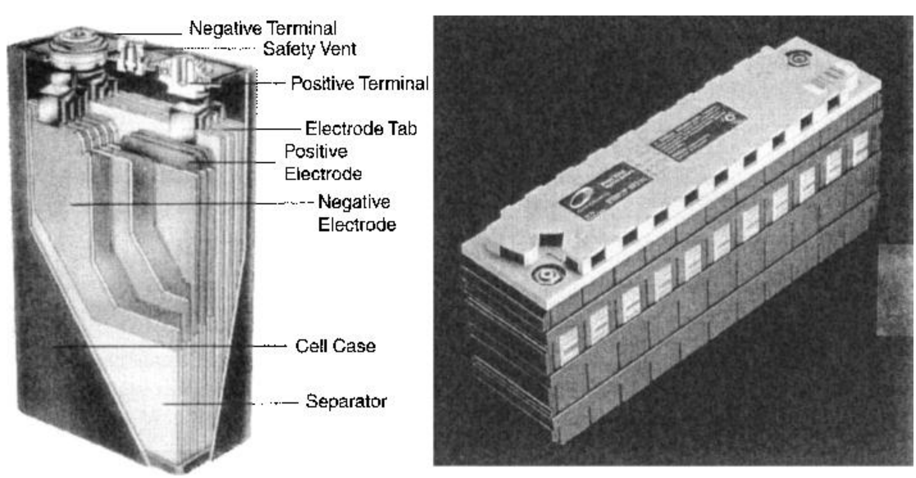
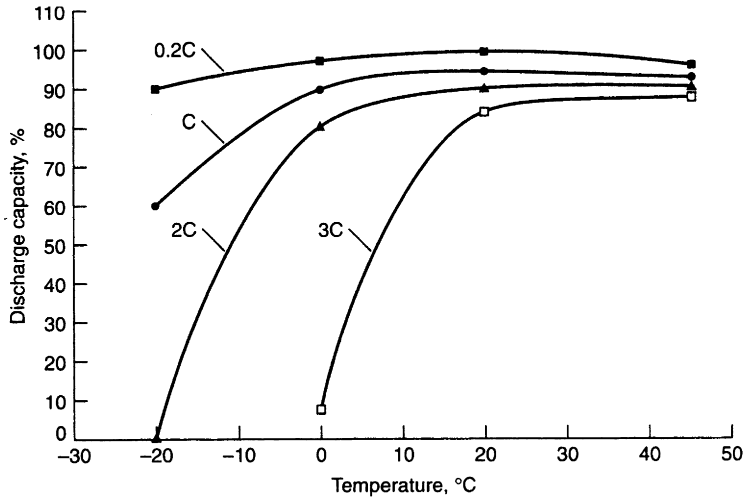
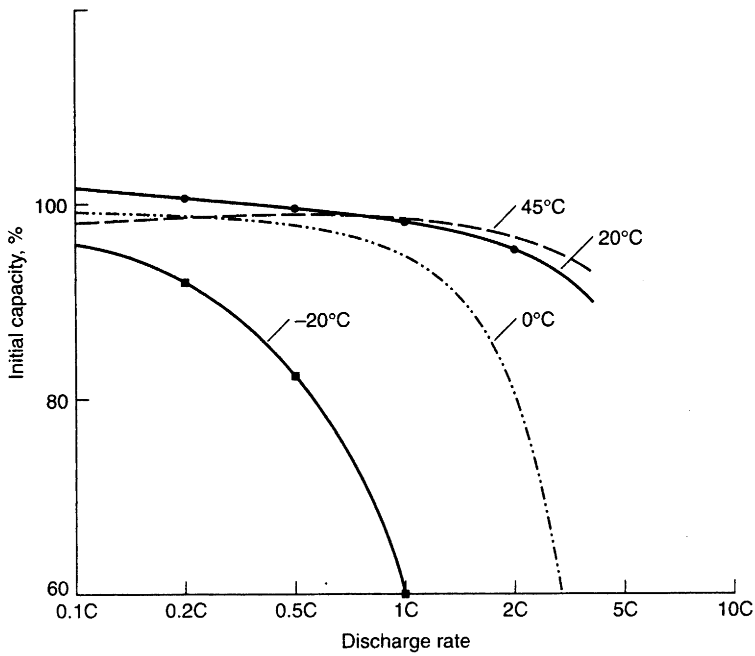


FIGURE 11-61 GM Ovonic 90-Ah prismatic cell and 13-V module.



(a)



(b)

FIGURE 11-62 (*a*) Discharge capacity vs. ambient temperature for sealed cylindrical nickel-metal hydride batteries at various discharge rates; end voltage 1.0 V/cell. (*b*) Discharge capacity (% of 0.2*C* rate) vs. discharge rate (*C*-rate) for sealed cylindrical nickel-metal hydride batteries at various temperatures: end voltage 1.0 V/cell.

memory effect, similar to nickel-cadmium cells, and this must be accommodated during use of the battery. However, there are several active research and development programs on the nickel-metal hydride technology, and improvements in performance and packaging are ongoing.

Rechargeable Ambient-Temperature Lithium Batteries. A huge range of new battery chemistries and products has become available in the last 5 to 10 years based on lithium. These batteries typically operate at close to room temperature, and provide significant advantages in gravimetric and volumetric energy and power characteristics relative to other rechargeable batteries. Products are available in an almost limitless range of electrochemistries and materials, and are known as lithium-ion, lithium polymer, lithium alloy, and lithium metal. The common material is lithium, but it may be present in metallic or ionic form, and there is a wide range of positive electrodes used, including cobalt oxide, manganese dioxide, and an equally wide range of electrolytes (e.g., organic liquids, gels, polymers), and conductive salts (e.g., lithium hexafluorophospate). It is impossible to generalize when describing this technology, but rather specific types and products must be considered. However, various versions of this technology are being used by the millions of units in many consumer electronics products, and they are being evaluated for larger industrial and propulsion applications.

Cell construction of a typical cylindrical lithium-ion cell is shown in Fig. 11-63. This technology is also available in prismatic and pouch cell designs. Cell capacities range from less than 1 Ah to over 150 Ah. A major advantage of the technology is the high operating voltage per cell, typically in the range of 2.5 to 4.2 V. This translates into fewer cells in series to achieve a given battery voltage. Typical performance data for small cylindrical cells as a function of discharge rate and temperature are shown in Fig. 11-64. Generally, the discharge power capability of this technology is not as good as that for nickel-cadmium batteries. These batteries are generally more expensive than commercial alternatives, but prices have been dropping as production capacity is increased.

It should be kept in mind that the use of rechargeable ambient-temperature lithium batteries must be carefully controlled to maintain battery safety. Many cells and battery packs include a controller which consists of electronics and software to carefully monitor, adjust, or terminate charge and discharge.

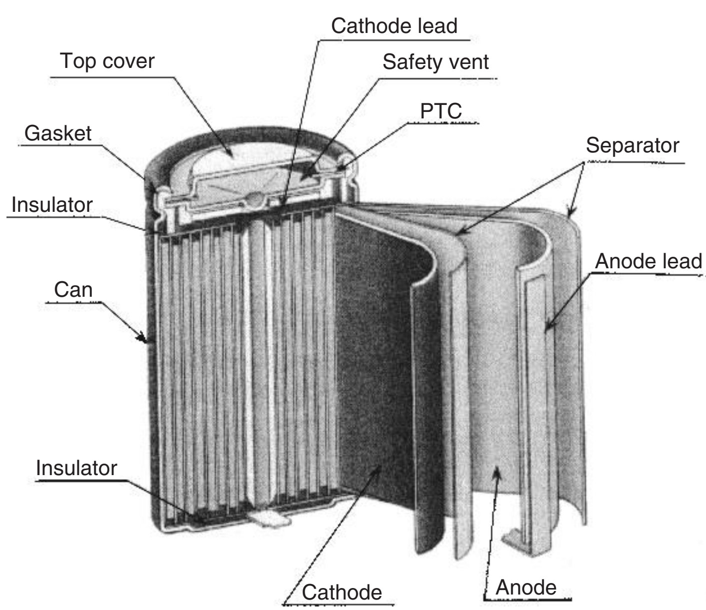


FIGURE 11-63 Cross-sectional view of a cylindrical Li-ion cell. (*Courtesy of the University of South Carolina. Reproduced with permission from the Journal of Power Sources.*)

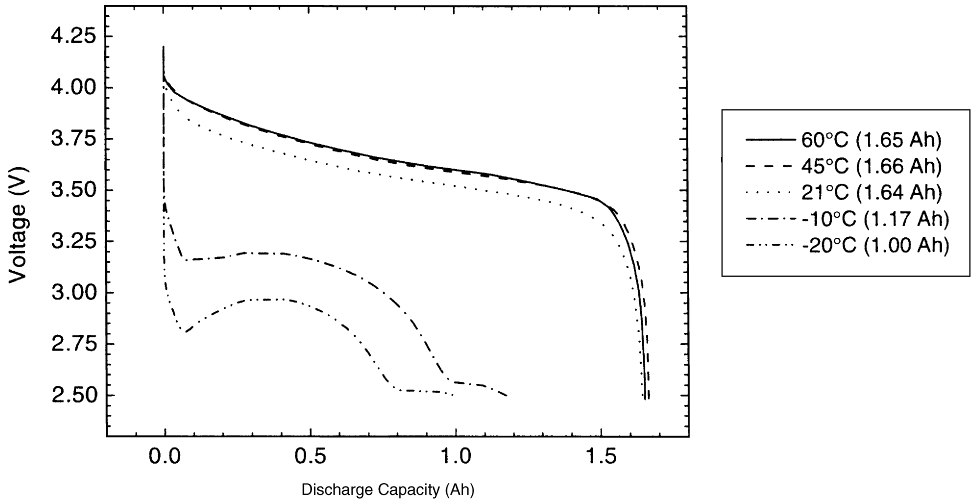


FIGURE 11-64 Approximate *C*-rate discharge of an 18650-type C/$LiCoO_2$ battery at various temperatures when charged in a CCCV regime at 1.65 A to 4.2 V for 2.5 hours, then discharged at 1.5 A. The average voltage at 21°C was 3.6 V, and at –20°C 2.9 V. (*Courtesy of NEC Moli Energy.*)

There have been numerous accidents involving this technology, and there are significant restrictions on transporting large quantities of cells and batteries on commercial aircraft, for example. This technology is also less tolerant to physical abuse than comparable battery chemistries, and care must be taken in harsh environments to avoid battery venting or thermal runaway.

The technology continues to evolve and improve in terms of performance and safety. There are many active research and development programs in progress worldwide to make products more robust, deliver higher discharge rates, and extend cycle life. Improvements in safety are also ongoing. This technology area is poised to continue growing for most small battery applications, and it may also become viable for large industrial and propulsion applications as further improvements in safety and control are realized.

The Battery Field. Readers are referred to an excellent in-depth technical review of batteries: *Modern Battery Technology,* Clive D. S. Tuck (ed.), West Sussex, England, Ellis Horwood Limited, ISBN 0-13-5902665-5. Another comprehensive battery reference is the *Handbook of Batteries,* 3rd ed., David Linden (ed.), New York, McGraw-Hill. This reference includes listings of companies that manufacturing various batteries.

The room-temperature solid electrolyte silver ion conducting devices described in association with Fig. 11-51 are available from Sigma Technologies International, Inc., Tucson, Arizona.

Several other advanced rechargeable battery technologies are being developed worldwide for many diverse applications. They are subject to government and industrial research and development, and the organizations and corporate sponsors/owners change frequently in this dynamic field. The following is a snapshot of the current advanced technologies in development.

- Lithium-alloy/iron sulfide high temperature batteries—Argonne National Laboratory
- Sodium/nickel chloride high temperature batteries (Zebra battery)—MES-DEA SA (Swiss), Beta Research and Development
- Sodium/sulfur high temperature batteries—NGK Insulator
- Vanadium-redox aqueous flowing electrolyte batteries—VRB Power Systems
- Zinc-air batteries—Electric Fuel Battery Corp
- Zinc/bromine flowing electrolyte batteries—ZBB Technologies

BIBLIOGRAPHY

Modern Battery Technology, Clive D. S. Tuck (ed.), West Sussex, England, Ellis Horwood Limited, ISBN 0-13-5902665-5.

Linden, D., and Reddy, T., *Handbook of Batteries*, 3rd ed., McGraw-Hill, New York, 2002.

11.9 FUEL CELLS

11.9.1 General Concepts

A *fuel cell* is an electrochemical device that continuously converts the chemical energy of a fuel (and oxidant) to electrical energy. The essential difference between a fuel cell and a battery is the continuous nature of the energy supply. The fuel and the oxidant, which is usually oxygen, are supplied continuously to a fuel cell from an external source.

The fuel cell uses liquid or gaseous fuels, such as hydrogen, hydrazine, hydrocarbons, and coal gas. The oxidant in a fuel cell is gaseous oxygen (or air). Fuel cells are close to receiving widespread commercial acceptance in stationary power generation due to improvements in efficiency and reductions in capital cost. Fuel cell-based generating stations typically have operating efficiencies ranging from 40% to 50%. In comparison, many generating stations operate at 30% to 35% efficiency.

A practical fuel cell power plant, depicted in Fig. 11-65, consists of at least three basic subsystems:

1. A power section, which consists of one or more fuel cell stacks–each stack containing many individual fuel cells usually connected in series to produce a stack output ranging from a few to several hundred volts (direct current). This section converts processed fuel and the oxidant into dc power.
2. A fuel subsystem that manages the fuel supply to the power section. This subsystem can range from simple flow controls to a complex fuel-processing facility. This subsystem processes fuel to the type required for use in the fuel cell (power section).
3. A power conditioner that converts the output from the power section to the type of power and quality required by the application. This subsystem could range from a simple voltage control to a sophisticated device that would convert the dc power to an ac power output.

In addition, a fuel cell power plant, depending on size, type, and sophistication, may require an oxidant subsystem as well as thermal and fluid management subsystems.

11.9.2 Operation of Fuel Cells

A simple fuel cell is illustrated in Fig. 11-66. Two catalyzed carbon electrodes are immersed in an electrolyte (acid in this illustration) and separated by a gas barrier. The fuel, in this case hydrogen, is bubbled across the surface of one electrode while the oxidant, in this case oxygen from ambient

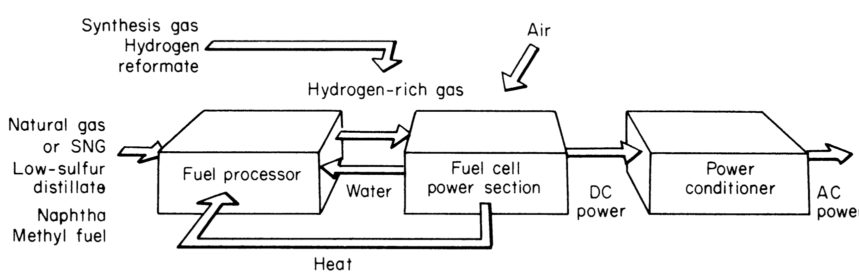


FIGURE 11-65 Generalized schematic diagram of a fuel cell power plant.

air, is bubbled across the other electrode. When the electrodes are electrically connected through an external load, the following events occur:

1. The hydrogen dissociates on the catalytic surface of the fuel electrode, forming hydrogen ions and electrons.
2. The hydrogen ions migrate through the electrolyte (and a gas barrier) to the catalytic surface of the oxygen electrode.
3. Simultaneously, the electrons move through the external circuit to the same catalytic surface.
4. The oxygen, hydrogen ions, and electrons combine on the oxygen electrode's catalytic surface to form water.

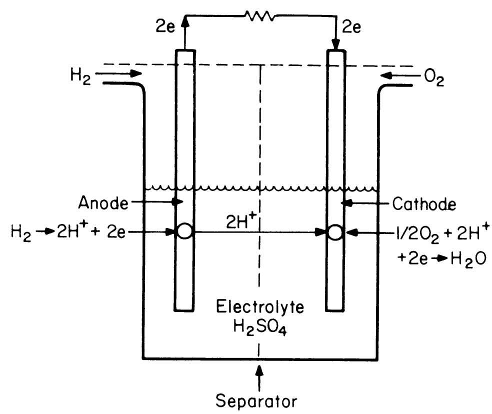


FIGURE 11-66 Operation (reaction mechanism) of the fuel cell.

The reaction mechanisms of this fuel cell, in acid and alkaline electrolytes, are shown in Table 11-15. The major differences, electrochemically, are that the ionic conductor in the acid electrolyte is the hydrogen ion (or, more correctly, the hydronium ion, H_3O^+) and the OH^- or hydroxyl ion in the alkaline electrolyte. Further, in the acid electrolyte the product, water, is produced as the cathode and in the alkaline electrolyte fuel cell at the anode.

The net reaction is that of hydrogen and oxygen producing water and electrical energy. As in the case of batteries, the reaction of one electrochemical equivalent of fuel theoretically will produce 26.8 Ah of dc electricity at a voltage that is a function of the free energy of fuel-oxidant reactions. At ambient conditions, this potential is ideally 1.23 V dc for a hydrogen-oxygen fuel cell.

11.9.3 Major Components of the Fuel Cell

The important components of the individual fuel cell are

1. The *anode* (fuel electrode) must provide a common interface for the fuel and electrolyte, catalyze the fuel oxidation reaction, and conduct electrons from the reaction site to the external circuit (or to a current collector that, in turn, conducts the electrons to the external circuit).
2. The *cathode* (oxygen electrode) must provide a common interface for the oxygen and the electrolyte, catalyze the oxygen reduction reaction, and conduct electrons from the external circuit to the oxygen electrode reaction site.
3. The *electrolyte* must transport one of the ionic species involved in the fuel and oxygen electrode reactions while preventing the conduction of electrons (electron conduction in the electrolyte causes a short circuit). In addition, in practical cells, the role of gas separation is usually provided by the electrolyte system. This is often accomplished by retaining the electrolyte in the pores of a matrix (or inert blotter). The capillary forces of the electrolyte within the pores allow the matrix to separate the gases, even under some pressure differential.

TABLE 11-15 Reaction Mechanisms of the H_2–O_2 Fuel Cell

	Acid electrolyte	Alkaline electrolyte
Anode	$H_2 \rightarrow 2H^+ + 2e$	$H_2 + 2OH^- \rightarrow 2H_2O + 2e$
Cathode	$\frac{1}{2}O_2 + 2H^+ + 2e \rightarrow H_2O$	$\frac{1}{2}O_2 + 2e + H_2O \rightarrow 2OH$
Overall	$H_2 + \frac{1}{2}O_2 \rightarrow H_2O$	$H_2 + \frac{1}{2}O_2 \rightarrow H_2O$

Other components also may be necessary to seal the cell, to provide for gas compartments, and to separate one cell from the next in a fuel cell stack.

11.9.4 General Performance Characteristics

The performance of a fuel cell is represented by the current density versus voltage (or polarization) curve (Fig. 11-67). Whereas ideally a single H_2-O_2 fuel cell could produce 1.23 V dc at ambient conditions, in practice, fuel cells produce useful voltage outputs that are somewhat less than the ideal and decrease with increasing load (current density). The losses or reductions in voltage from the ideal are referred to as *polarization*, as illustrated in Fig. 11-67.

These losses include the following:

1. Activation polarization represents energy losses that are associated with the electrode reactions.
2. Ohmic polarization represents the summation of all the ohmic losses within the cell, including electronic impedances through electrodes, contacts, and current collectors and ionic impedance through the electrolyte. These losses follow Ohm's law.
3. Concentration polarization represents the energy losses associated with mass transport effects. For instance, the performance of an electrode reaction may be inhibited by the inability of reactants to diffuse to or products to diffuse away from the reaction site.

The net result of these polarizations is that practical fuel cells produce between 0.5 and 0.9 V dc at currents of 100 to 400 mA/cm^2 of cell area. Fuel cell performances can be increased by increasing cell temperature and reactant partial pressure. For any fuel cell, the tradeoff always exists between achieving higher performance by operating at higher temperature or pressure and confronting the materials and hardware problems imposed at the more severe conditions.

11.9.5 Fuel Cell Systems

Classification and Types. Fuel cell systems can take a number of different configurations, depending on the combination of type of fuel and oxidant, whether the fueling is direct or indirect, the type

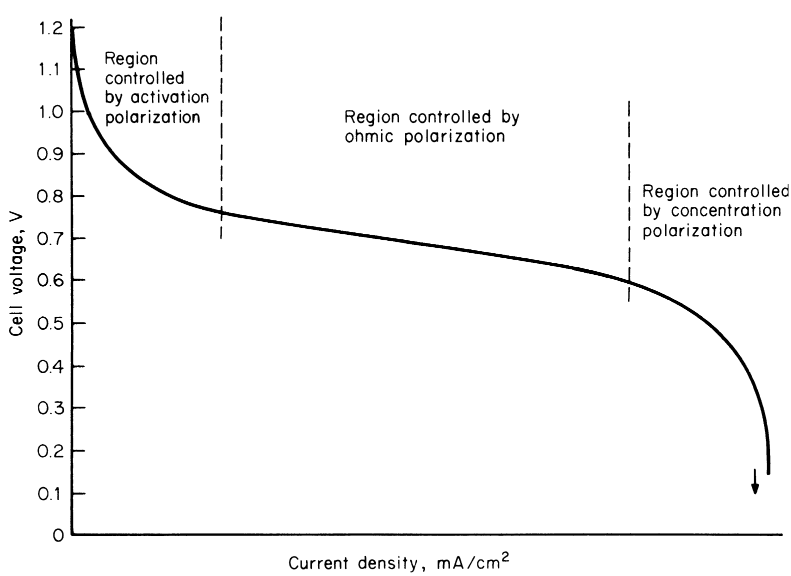


FIGURE 11-67 Fuel cell polarization curve.

of electrolyte, the temperature of operation, etc., although in actual practice, the number of combinations is limited. A listing of the practical fuel cell systems is given in Table 11-16.

Acid Fuel Cells. Table 11-16 lists two types of acid fuel cells (solid polymer electrolyte and phosphoric acid), aqueous alkaline fuel cells, molten carbonate fuel cells, and solid oxide fuel cells. Acid fuel cells are characterized by the following:

Ionic conduction is provided by hydrogen ions [or by hydronium ions (H_3O^+)].

Platinum or platinum alloys (in very small quantity) are the active electrocatalysts.

Carbon (graphite) is an acceptable material of construction for current collectors, gas separators, etc., and is commonly used.

Solid Polymer Electrolyte System. The solid polymer electrolyte (SPE) system uses an ion-exchange membrane as the electrolyte. The advantages of the SPE fuel cell are (1) the electrolyte, being a solid, cannot change, move about, or vaporize from the system; and (2) the only liquid in the fuel cell is water, minimizing corrosion. The disadvantages are (1) the SPE must be hydrated (water-saturated) to perform; consequently, operation must be under conditions where the by-product water does not vaporize into the reaction air stream faster than it is produced; this constrains cell operation to under 60°C at ambient pressure and about 120°C at elevated pressures; and (2) the SPE freezes at about 0°C and undergoes a freeze-drying phenomenon. This constraint is applicable to those where low-temperature capability is not a requirement.

Due to their inability to operate much above 120°C, SPE fuel cells are best suited for use with hydrogen-rich gases that contain little or no carbon monoxide. Carbon monoxide inhibits the fuel cell anode reaction, the degree of inhibition decreasing with increasing temperature. Consequently, SPE fuel cells have found their important applications in the space program operating on pure hydrogen or in military applications operating on hydrogen obtained by the decomposition of a hydride.

TABLE 11-16 Classification of Practical Fuel Cells

Application	Fuel	Oxidant	Electrolyte	Temperature
Remote				
Space	Direct H_2	Liquid O_2	Aqueous alkaline Solid polymer	Low, intermediate Low
Undersea	Direct H_2 Direct hydrazine	Liquid O_2 Hydrogen peroxide	Aqueous alkaline	Low
Military				
Low power, ≤ 100 W	Indirect hydride	Air	Aqueous alkaline Solid polymer	Low
High power, ≥ 500 W	Indirect hydrocarbon Indirect methanol	Air	Phosphoric acid	Intermediate
Commercial power				
Dispersed (or on-site)	Indirect hydrocarbon		Phosphoric acid	Intermediate
	Indirect methanol, ethanol Direct coal gas	Air	Molten carbonate	High
Central station	Indirect coal Direct coal gas	Air	Phosphoric acid Molten carbonate Solid oxide	Intermediate High Very high
Vehicle	Hydrogen Hydride Indirect methanol Indirect hydrocarbon	Air	Phosphoric acid	Intermediate

Phosphoric Acid Electrolyte System. The phosphoric acid electrolyte system operates at 150 to 220°C. At lower temperatures, phosphoric acid is a poor ionic conductor. At higher temperatures, material stability (carbon and platinum) becomes limiting. The advantages of phosphoric acid fuel cells are (1) the electrolyte is very stable, (2) the phosphoric acid can be highly concentrated (~ 100%) where the water vapor pressure is very low and steady-state water removal by the reactant gases will always equal product water rate, and (3) at 150 to 220°C, the anode performance is very good even on fuels containing up to 5% carbon monoxide. The disadvantage of phosphoric acid fuel cells is that the cathode performance is sluggish. In fact, the major technology thrusts in phosphoric acid are toward improvement of the cathode. The phosphoric acid fuel cell is a preferred system for use with fuels containing carbon oxides.

Alkaline Fuel Cells. Although early alkaline fuel cells operated at relatively high temperature (~250°C) with concentrated (85 wt%) potassium hydroxide, systems developed more recently operate at much lower temperatures (<120°C) using less concentrated (35 to 50 wt%) potassium hydroxide. The lower temperature enables the use of matrices to retain the electrolyte and increases the life of other components.

Alkaline fuel cells are characterized by the following:

Ionic conduction is provided by hydroxyl (OH^-) ions.

A wide range of electrocatalysts can be used, including nickel, silver, metal oxides, spinels, and noble metals—although the truly high-performance systems use at least small amounts of noble metal.

Construction materials include carbon, nickel, and stainless steel.

The advantages of alkaline fuel cells are that (1) cathode performance is much better than for acid fuel cells and (2) materials of construction tend to be low in cost. The primary disadvantage is that the electrolyte reacts with carbon oxides to form potassium carbonate. This severely limits the cells' performance. Thus, alkaline fuel cells have only limited application where carbonaceous fuels or air is used as a reactant. The important applications (space and underseas) involve pure hydrogen and oxygen.

Molten Carbonate Fuel Cells. Molten carbonate fuel cells use an alkali metal (Li, K, Na) carbonate as the electrolyte. Since these salts can function as electrolytes only when in the liquid phase, the cells operate at 600 to 700°C, which is above the melting points of the respective carbonates. Molten carbonate cells are characterized by the following:

Ionic conduction is by the carbonate ion; thus, the carbonate ion must be involved in the two electrode reactions:

$$\begin{array}{ll} {}^1/_2\,O_2 + CO_2 + 2e \rightarrow CO_3^{2-} & \text{(cathode)} \\ CO_3^{2-} + H_2 \rightarrow CO_2 + H_2O + 2e & \text{(anode)} \\ \hline {}^1/_2\,O_2 + H_2 \rightarrow H_2O & \text{(net)} \end{array}$$

A consequence of this is that CO_2 must be recycled from the anode to the cathode.

At 600 to 700°C, electrode reactions proceed without highly specific catalysts. Nickel and nickel oxide work quite well; noble metals are not used.

Construction materials include nickel and ceramics.

The advantages of molten carbonate fuel cells are that (1) cell performance is good, activation polarization is small; (2) at 600 to 700°C, any carbon monoxide in the fuel converts to hydrogen on the anode via the water gas shift reaction, that is,

$$CO + H_2O \rightarrow CO_2 + H_2$$

(as a result, fuel gases high in carbon monoxide are readily used); and (3) waste heat from the fuel cell can be available at a relatively high temperature (>500°C), enabling its use in bottoming or industrial heating cycles.

Disadvantages are that (1) the high temperature imposes severe constraints on materials suitable for long lifetimes and (2) a course of carbon dioxide is required to complete the cathode reaction (this is provided by recycling CO_2 from the anode exhaust to the cathode inlet). As a result, molten carbonate fuel cells are best suited for applications that integrate the fuel cell with a carbonaceous fuel processor, that is, a reformer or coal gasifier.

Solid Oxide Fuel Cells. As the name implies, solid oxide fuel cells employ a solid, nonporous metal oxide electrolyte, which allows ionic conductivity by the migration of oxide ions through the lattice of the crystal. Stabilized zirconia is commonly used as the electrolyte. The cells operate at 900 to 1,000°C. Whereas practical cells of the technologies previously discussed are normally packaged into "filter press" or "plate and frame" stack assemblies, solid oxide fuel cells are configured into tubular cell stacks.

Characteristics of the solid oxide fuel cell include the following:

Ionic conduction is provided by oxide ions.

The cathode employs metal oxides, such as praseodymium oxide or indium oxide; the anode uses nickel or nickel cermet.

Because of the high temperature, materials of construction will likely be confined to ceramics or metal oxides.

Solid oxide fuel cells offer advantages similar to those of molten carbonate cells, that is, good performance on fuels containing hydrogen or hydrogen and carbon monoxide, the elimination of noble-metal catalysts, and the availability of high-grade reject heat. In addition, they do not suffer the constraint of molten carbonate cells which require a carbon dioxide recycle to the cathode. The primary disadvantages are the very high temperature of operation and the severe material constraints imposed by the ~1,000°C temperature.

11.9.6 Low-Power Fuel Cell Systems

The advantageous characteristics of fuel cells led to the development of a number of different systems ranging in size from the portable units, 5 W or smaller to kilowatt power levels (where ease of operation, low maintenance, and silence are important), to large stationary plants delivering megawatts of power (where the high efficiency over the range from full to partial load and reduced pollution are significant). The lower-power fuel cells were designed mainly for military or special applications such as the space program. For the space applications and for forward-area military use, the fuel cell offers high energy densities that exceed the performance of batteries when operated over long periods of time.

Portable Fuel Cells. Fuel cells in the power range of up to about 200 W can be an attractive alternative to batteries for long-term operation, with potential reduction in the weight and cost. The size limitation of these portable fuel cell systems is generally too small to permit the use of elaborate fuel conditioning. Hence, easily handled and readily oxidized fuels (e.g., liquid or gaseous fuels such as hydrazine, methanol, and ammonia) and fuels that provide hydrogen through a simple physical or chemical reaction have been used in most fuel cell systems of this size and type. For these ground applications, air-breathing, rather than pure oxygen, systems are used.

While the fuel cell systems in this size and power range potentially have advantageous characteristics, system complexities and the development of new battery systems with higher energy densities have limited the interest and successful use of the low-power fuel cell.

Direct Fuel Cell Systems. Direct-type fuel cells, in which the fuel can be introduced into the fuel cell without requiring conversion to hydrogen, were considered for small fuel cell systems because they eliminated the need for a fuel-conditioning unit, thus saving important space and weight. Methanol (CH_3OH) and hydrazine (N_2H_4) were the main liquid fuels used. Methanol is directly oxidizable, but removal of the carbonate, one of the reaction products of the dissolved methanol fuel cell in alkaline electrolytes, from the electrolyte is extremely difficult. Efforts then shifted to hydrazine. Hydrazine decomposes easily into hydrogen and nitrogen at the electrode surface; in fact, the voltage observed is that of hydrogen.

Major effort was directed toward a silent power source for forward-area military use. A 60-W, 24-V hydrazine-air fuel cell was developed in a configuration similar to the one used later for the metal hydride cell. The fuel cell used a 35% potassium hydroxide electrolyte and a 64% hydrazine monohydrate fuel and operated between 55 and 70°C with a fuel utilization of 600 Wh/kg. A larger 300-W, 24-V power source also was developed for forward-area use. This system weighed 20 kg, with electrolyte and 4 L of fuel and had a volume of 35 dm^3. The fuel was sufficient for 12 h of operation at 300 W. Field tests confirmed the successful electrochemical functioning of the cell, but mechanical deficiencies caused early failure of the system.

Metal Hydride Fuel Cells. The majority of current portable fuel cell developments use a metal hydride as the source of hydrogen fuel. Metal hydrides are attractive because they can store large amounts of hydrogen more conveniently and with a higher energy density (total equivalents of hydrogen per total weight of hydrogen source and container) than hydrogen in a pressurized or liquefied form.

One type of metal hydride produces hydrogen by the reaction with water, for example, calcium hydride (CaH_2):

$$CaH_2 + 2H_2O \rightarrow Ca(OH)_2 + H_2$$

A second type of metal hydride, a reversible hydride, is based on the principle that certain metals or alloys (e.g., iron titanium, lanthanum nickel, and various other rare-earth metal and nickel alloys) have the ability to take up large amounts of hydrogen gas within their crystal structure. A reduction in pressure or an increase in temperature release the hydrogen. These hydrides can deliver hydrogen at about 500 Wh/kg.

$$NaAlH_4 + H_2O \text{ (excess)} \rightarrow Al(OH)_3 + NaOH + 4H_2$$

The hydride, in the form of a solid pellet, delivers in excess of 2000 Wh/kg. Hydrogen for 4 h of operation is supplied to the fuel cell with a single 120-g charge. The Kipp generator delivers hydrogen to the fuel cell on demand; when no hydrogen demand exists, the pressure builds up in the generator, forcing water away from the fuel pellet and stopping the reaction.

A later design used a reversible metal hydride, lanthanum pentanickel hydride, as the source of hydrogen:

$$LaNi_5 + 3H_2 \rightleftharpoons LaNi_5H_4$$

This change, replacing the exothermic sodium aluminum hydride generator with the reversible metal hydride source, which absorbs heat on the release of hydrogen, could reduce the total heat output of the system by 65%. The system was found to operate satisfactorily at 20°C, but higher ambient temperatures still caused problems because of the much higher operating temperature of the fuel cell stack.

SPE Fuel Cell. There is renewed interest in the SPE fuel cell for applications in which mobility is important and power requirements are low. The advantages of the SPE cell are ease of product water removal, simple construction, stable electrode-electrolyte interface, and favorable life characteristics.

The SPE cell is being considered for power levels from a few watts to 500 W, operating in the –40 to 50°C range with no restrictions on humidity. Special designs, including insulation for low temperatures and methods for waste heat disposal, probably will be required to achieve this performance. The cell operates on hydrogen and ambient air. Preferred fuel sources for hydrogen generation

TABLE 11-17 Utility Fuel Cell Power Plant Programs

Role	Size	Fuel	Efficiency, %	Electrolyte
On-site power plants	40–300 kW	Pipeline gas	39–42 or 90 (with reject-heat recovery)	Phosphoric acid
Dispersed (substation) power plants	5–25 MW	Petroleum- or coal-derived gas or liquid	41–47 or 80 (with reject-heat recovery)	Phosphoric acid
Central station power plants	150–600 MW	Coal	45–50	Molten carbonate or phosphoric acid

are magnesium and aluminum, which are reacted with salt water. Bottled hydrogen, hydride-stored hydrogen, or hydrides also may be used. With magnesium, it is expected to obtain energy densities (fuel consumption) of about 220 Wh/kg on a wet basis and 1300 Wh/kg on a dry basis. For longer missions, where a larger system weight can be tolerated, re-formed methanol combined with carbon monoxide absorption is being considered.

Utility Fuel Cell Power Plants. Interest in the fuel cell as a utility power plant results from its efficiency, its environmental acceptability, and its modular construction. The fuel cell may serve utilities in several ways, as summarized in Table 11-17. Relatively small fuel cell power plants (with a capacity ranging from 40 to 300 kW) could be set up in commercial and residential buildings. Such a plant would use natural gas as a fuel. The plant would provide both electric and thermal energy (the latter from the waste heat of the fuel cells), consuming the same amount of fuel ordinarily required for the thermal demand alone. Overall efficiencies approaching 90% have been projected for fuel cell power plants of this type. An advantage of the fuel cell is that the waste heat can be used without altering the power production characteristics. Larger plants, ranging in capacity from 5 to 25 MW, could be dispersed throughout an electric utility system to perform load-following duty efficiently, taking advantage of the higher efficiency of the fuel cell even at reduced loads. In the future, fuel cells could be integrated with coal gasifiers to provide large, central-station, base-load power plants that utilize coal directly. The capacity of such plants would range from 150 to 1,000 MW. A plant of this kind is projected to be more than 45% efficient, on the basis of the heating value of the coal consumed.

Phosphoric Acid Fuel Cells. The basic *cell structure* of the phosphoric acid fuel cell (PAFC) is shown in Fig. 11-68. It consists of

1. *A carbon or graphite separator-current collector plate* that separates hydrogen from the air of the adjacent cell (in a multicell stack) and also provides the electrical series connection between cells.
2. *Anode current collector ribs* that conduct the electrons from the anode to the separator plate. The ribbed configuration provides gas passages for hydrogen distribution to the anode.
3. *An anode* that consists of a porous graphitic substrate with the surface adjacent to the electrolyte treated with a platinum or platinum alloy catalyst.
4. *An electrolyte matrix* that retains the concentrated phosphoric acid.
5. *A cathode* that is similar to the anode but uses a modified noble-metal catalyst and an increased catalyst loading (usually 0.5 mg/cm^2) to enhance the oxygen reduction kinetics.
6. *Cathode current collector ribs* that are also virtually identical to the anode ribs.

These single cells are stacked in series to produce the desired output power and voltage.

Phosphoric acid fuel cell power plants are also being considered for use in central stations. These power plants would utilize integrated coal gasifiers to supply the hydrogen-rich fuel and would provide base load power.

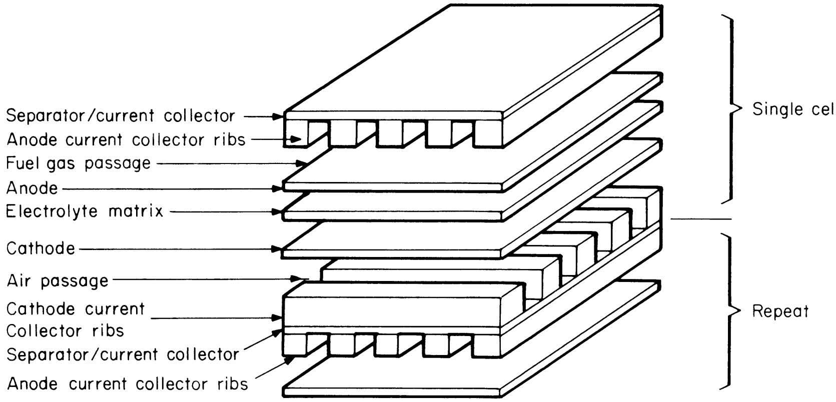


FIGURE 11-68 Basic phosphoric acid electrolyte fuel cell.

Molten Carbonate Fuel Cells. In contrast to the PAFC, molten carbon fuel cell (MCFC) development efforts are focused on large, multimegawatt (~600 MW) central-station power plants integrated with a coal gasifier. This focus is, in part, due to their high operating temperature and the problems inherent in starting (heating) and stopping (cooling) MCFC power plants.

Cell structure for the MCFC is geometrically very similar to that of the PAFC. The materials that are used, however, are very different from those used in the PAFC. The MCFC consists of

1. A *separator and current-collector plate* that separates the fuel gas from the air of the adjacent cell in a multicell stack and also provides the electrical connection between cells. Like its PAFC counterpart, it must be impermeable to hydrogen and oxygen, a good electronic conductor, and stable to fuel and air environments in the presence of 650°C carbonate salts.
2. An *anode current collector* that conducts the electrons from the anode to the separator plate. This current collector also must provide passage for fuel flow. In some configurations, this function is provided by ribbing or folding the separator plate.
3. An *anode* that consists of a porous nickel treated with a refractory oxide to reduce sintering. At the 650°C temperature, no other catalyst is required.
4. An *electrolyte system* comprising a mixture of lithium–potassium carbonate and inert powder (presently a lithium aluminate). This mixture forms a paste when molten and freezes to form a "tile" when cooled. This electrolyte system presents major challenges to MCFC development. It must have minimum ionic resistance while separating the fuel and oxidant gases at pressure differentials in excess of 0.07×10^5 Pa. In addition, it must be electronically insulating.
5. A *cathode* that is similar to the anode except that it uses nickel oxide (doped with lithium to impart electronic conductivity).
6. A *cathode current collector* that has similar requirements and configurational operations as the anode current collector. Since nickel is thermodynamically unstable, material options include lithium-doped nickel oxide and corrosion-resistant stainless steel.

As with the PAFC, individual cells are stacked in series to result in a cell stack of the required power and voltage output.

Molten Carbonate Fuel Cell Power Plants. The thrust of the MCFC program is the development of a central-station power plant, comprising a coal gasifier, gas cleanup system, MCFC

TABLE 11-18 Design Requirements and Goals for a Central Station MCFC Power Plant

Requirements	
Central station plant	
Power level	–675 MW(e)
Fuel specification	Illinois no. 6 coal
Modular construction	
Environmental	Projected 1985 federal requirements
Site characteristics	"Middletown" except for cooling tower heat rejection
Goals	
Base load duty with daily load following capability	
Heat rate	6.8×10^6 J/kWh
Capital installed cost (1982 dollars)	\$1500/kW(e)
Plant availability	85%
Life goals (75% capacity factor)	
Fuel cell stacks	6 years
Balance of plant	30 years
Startup/shutdown	
Startup: Cold startup in 4 to 6 h	
Shutdown: 100% to zero load in 3 h	
Daily load following	
Large-load-change response time of 2 h	
Small-load-change response rate up to 2%/min	
Abnormal conditions	
Complete-load rejection (breakers opening)	
Partial-load rejection (from power system breakup)	
Sustained abnormal voltage or frequency operation	
Limit fault current to 1.1 per unit current (rms basis)	
Other	
Independent var control	

(topping cycle), and gas or steam turbine (bottoming cycle). Design requirements and goals for a 675-MW power plant are described in Table 11-18. A possible power plant configuration would include

600 MCFC stacks (having five hundred 1-m^2 cells each) for a total 450-MW output

15 coal gasifiers capable of handling 3×10^{11} J/h each

10 heat-recovery steam generators

Five 15-MW gas turbines

One 150-MW steam turbine

Although the ultimate application of the MCFC will likely be as a large coal-fueled, central-station power plant, other applications such as that of dispersed or on-site generators with and without reject heat recovery are also being considered. Because of the high operating temperature, the quality of the MCFC's waste heat could be compatible with a variety of industrial heating applications. Also, at 600°C, in situ re-forming of methane or methanol is possible; this could result in a very efficient small power plant.

Direct Fuel Cell Systems. Methanol (CH_3OH) and hydrazine (N_2H_4) are the liquid fuels that have been used in direct fuel cells, since they are more readily oxidized than the fossil fuels. Work is still in the development stage, and no systems were available commercially in the mid-1970s. Effort in recent years has been deemphasized.

Hydrazine has been used in a 60-W and larger configurations. The 60-W hydrazine unit is similar to the hydrogen fuel cell and delivers over 600 Wh/kg but was abandoned in favor of the hydrogen system. Similar experience was obtained with a larger 120-kg, 1.5-kW unit which degraded rapidly after 300 h of service.

A relatively small effort continues on direct hydrocarbon fuel cells, particularly with high-temperature (1,000 to 1,200°C) solid electrolyte fuel cells. The higher temperatures should allow direct reaction of the hydrocarbons with improved kinetics, although the potential benefits may be offset by the problems of high-temperature operation.

11.9.7 Fuel Cell Resources

Web sites

http://www.fuelcells.org

General Motors, http://www.gm.com

Fuel cell Today, http://www.fuelcelltoday.com

United Technologies, http://www.utcpower.com

11.10 MAGNETOHYDRODYNAMICS

By N.B. MORLEY and M.S. TILLACK

11.10.1 Introduction

The interaction of moving conducting fluids with electric and magnetic fields provides for a rich variety of phenomena associated with electro-fluid-mechanical energy conversion. Effects from such interactions can be observed in liquids, gases, two-phase mixtures, or plasmas. Numerous scientific and technical applications exist, such as heating and flow control in metals processing, power generation from conducting two-phase mixtures or conductor-seeded high-temperature gases, magnetic confinement of high-temperature plasmas—even dynamos that create magnetic fields in planetary bodies. Several terms have been applied to the broad field of electromagnetic effects in conducting fluids, such as *magneto-fluid-mechanics, magneto-gas-dynamics*, and the more common one used here—*magnetohydrodynamics*, or MHD.

Practical MHD devices have been in use since the early part of the twentieth century. For example, an MHD pump prototype was built as early as 1907.[1] More recently, MHD devices have been used for stirring, levitating, and otherwise controlling flows of liquid metals for metallurgical processing and other applications.[2,3] Gas-phase MHD is probably best known in MHD power generation. Since 1959,[4,5] major efforts have been carried out around the world to develop this technology in order to improve electric conversion efficiency, increase reliability by eliminating moving parts, and reduce emissions from coal and gas plants. Closed-cycle liquid metal MHD systems using both single-phase and two-phase flows also have been explored.

Still, more novel applications are in development or on the horizon. For example, recent research has shown the possibility of seawater propulsion using MHD[6] and control of turbulent boundary layers to reduce drag.[7] Extensive worldwide research on magnetic confinement of plasmas has led to attainment of conditions approaching those needed to sustain fusion reactions.[8]

In the following sections, we review the basic equations describing coupled MHD behavior as well as some basic MHD phenomena in liquids, gases, and two-phase mixtures. Much of the underlying physics described is common to many of the applications cited above. Also included are discussions of several of the most important applications, together with their special analysis techniques and examples of equipment involved.

11.10.2 Basic Equations

The Full Set of MHD Equations

The MHD Equations. The complete set of MHD equations for a Newtonian, constant-property fluid flow includes the Navier-Stokes equations of motion (i.e., momentum equation), the equation of mass continuity, Maxwell's equations, and Ohm's law. In differential form, they constitute the following system of equations:

$$\rho\left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\,\mathbf{u}\right) = -\nabla p + \mathbf{j} \times \mathbf{B} + \mu_f \nabla^2 u + \rho \mathbf{g} \tag{11-14}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \tag{11-15}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{11-16}$$

$$\nabla \times \mathbf{B} = \mu_m \mathbf{j} \tag{11-17}$$

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{11-18}$$

where the MHD body force $\mathbf{j} \times \mathbf{B}$ is included in the Navier-Stokes equation. The displacement current has been neglected from Ampere's law, which is a valid approximation for nonrelativistic phenomena typical of the response of an inertial liquid. Implicit in Eqs. (11-14) to (11-18) are the following additional relations:

$$\nabla \cdot \mathbf{B} = 0 \tag{11-19}$$

$$\nabla \cdot \mathbf{j} = 0 \tag{11-20}$$

This system of equations is a rich one, describing not only all of the phenomena generally associated with low frequency fluid mechanics and electromagnetics, but also new phenomena not seen in either discipline. Simplifications usually are required to obtain solutions to physical systems of interest. For instance, for quasi-steady flow problems where $\partial B/\partial t$ is negligible, the electric field can be represented as the gradient of an electric potential φ, which simplifies the problem by eliminating vector Eq. (11-16).

Magnetic Induction. The magnetic induction equation is derived easily by taking the curl of Ohm's Law:

$$\nabla \times \mathbf{j}/\sigma = \nabla \times \mathbf{E} + \nabla \times (\mathbf{u} \times \mathbf{B}) \tag{11-21}$$

If $\nabla \times \mathbf{E}$ is replaced by Faraday's Law (Eq. 11-16) and $\nabla \times \mathbf{j}$ is replaced by the curl of Ampere's law (Eq. 11-17), then, using the vector identity

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} \tag{11-22}$$

we obtain

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \frac{1}{\mu_m \sigma} \nabla^2 \mathbf{B} \tag{11-23}$$

Equation (11-23) is known as the induction equation, and suggests that the motion of a conducting liquid in an applied magnetic field, through the generation of electric current, will induce a magnetic field in the medium. The total field is the sum of the applied and induced magnetic fields. The relative strength of the induced field is characterized by the magnetic Reynolds number ($\mathrm{Re}_m = \sigma\mu_m uL$). The neglect of the induced magnetic field is a valid assumption when Re_m is small.

Dimensionless Parameters. Fluid mechanics equations typically are cast in dimensionless form so that the relative strengths of the different terms can be inferred by the size of any multiplying factors. The equation of motion (Eq. 11-14) can be written in dimensionless form by making the substitutions

$$\mathbf{j}^* = \frac{\mathbf{j}}{\sigma u_o B_o} \tag{11-24}$$

$$p^* = \frac{p}{\sigma u_o B_o^2 a} \tag{11-25}$$

$$\nabla^* = \frac{1}{a}\nabla \qquad u^* = u/u_o \qquad B^* = B/B_o \tag{11-26}$$

where a, u_o, and B_o are characteristic values of length, velocity, and applied magnetic field. Characteristic values of the current density and pressure have been selected carefully in order to scale the phenomena of interest; different values could have been selected, leading to different systems of nondimensionalization. Using this system, the equation of motion (excluding gravity) becomes

$$\frac{1}{N}\left(\frac{\partial \mathbf{u}^*}{\partial t} + (\mathbf{u}^* \cdot \nabla)\mathbf{u}^*\right) = -\nabla p^* + \mathbf{j}^* \times \mathbf{B}^* + \frac{1}{Ha^2}\nabla^2\mathbf{u}^* \tag{11-27}$$

The characteristic parameters Re, Ha, and N are the Reynolds number, the Hartmann number (which is an average measure of the ratio of magnetic to viscous force), and the interaction parameter (which is a measure of the ratio of magnetic to inertial forces). They are defined as

$$\mathrm{Re} = \rho u_o a/\mu_f \tag{11-28}$$

$$\mathrm{Ha} = aB_o\sqrt{\sigma_f/\mu_f} \tag{11-29}$$

$$N = \mathrm{Ha}^2/\mathrm{Re} = aB_o^2\sigma_f/\rho u_o \tag{11-30}$$

Table 11-19 gives representative values of these characteristic dimensionless parameters, for example, cases of interest. When the Hartmann number and interaction parameter are both sufficiently large, the momentum equation (Eq. 11-27) throughout the bulk of the fluid is often reduced to the simple form

$$\nabla p = \mathbf{j} \times \mathbf{B} \tag{11.31}$$

Whether this inviscid, intertialess approximation is always valid in cases with high Hartmann number and interaction parameter is still the subject of some debate.

Electrical Equations and Ohm's Law.

The Lorentz Force. Underlying the MHD body force is the fact that free charges, with charge q, moving in a magnetic field experience a "Lorentz" force perpendicular to both their velocity and the magnetic field induction:

$$\mathbf{F}_q = q(\mathbf{v} \times \mathbf{B}) \tag{11-32}$$

TABLE 11-19 Typical Values of Re, Ha, N, and Re_m for Several Materials (Assuming $a = 0.1$ m, $B = 1$ T, $u = 1$ m/s)

	NaK (100°C)	Hg (20°C)	Electrolyte (20°C) 15% KOH	Air (3,000°C) with 2% K
Re	1.6×10^5	9.1×10^5	4.3×10^4	350
Ha	6800	2700	17.5	98
N	290	8.2	7×10^{-3}	27
Re_m	0.30	0.14	1.2×10^{-5}	1.3×10^{-5}

For collisionless particles of mass m, the Lorentz force results in pure harmonic motions in the plane perpendicular to the magnetic field (B_z) with characteristic cyclotron frequency $\omega_c = qB_z/m$:

$$m\dot{v}_y = -qv_xB_z \qquad m\dot{v}_x = qv_yB_z \tag{11-33}$$

$$\ddot{v}_y = -\frac{qB_z}{m}\dot{v}_x = -\left(\frac{qB_z}{m}\right)^2 v_y \qquad \ddot{v}_x = \frac{qB_z}{m}\dot{v}_y = -\left(\frac{qB_z}{m}\right)^2 v_x \tag{11-34}$$

$$v_y = A_1 \cos \omega_c t + A_2 \sin \omega_c t \qquad v_x = A_3 \cos \omega_c t + A_4 \sin \omega_c t \tag{11-35}$$

In contrast, for collisional particles that are forced to follow the fluid velocity **u**, the Lorentz force acts on electrons and ions in a direction perpendicular to the flow, but in opposite directions for positive and negative charges. The net result is charge separation, leading to the generation of electric fields. The open circuit voltage between electrodes spaced a distance d apart in a conducting fluid is

$$V_{oc} = \int_0^d (\mathbf{u} \times \mathbf{B}) \cdot dI \tag{11-36}$$

An electric field $\boldsymbol{\varepsilon}$ arises between the electrodes such that $\boldsymbol{\varepsilon} + \mathbf{u} \times \mathbf{B} = 0$, corresponding to the zero-current condition in Ohm's Law (Eq. 11-18). If current is allowed to flow, as a result of some return current path, then the electric field and electrode voltage are reduced due to the electrical resistance of the fluid:

$$\boldsymbol{\varepsilon} = (\mathbf{u} \times \mathbf{B} - \mathbf{j}/\sigma) \tag{11-37}$$

The Hall Effect. In the regime between collisional and collisionless particles, the Hall effect can be important. Usually, the current induced in the fluid is carried predominantly by electrons, which are considerably more mobile than ions. The electron drift velocity, given by

$$\mathbf{j} = n_e e\, \mathbf{u}_e \tag{11-38}$$

leads to a second component of velocity, and so, according to Eq. (11-32), a secondary force and electric field

$$\varepsilon H = \beta \mathbf{j} \times \mathbf{B} \tag{11-39}$$

where $\beta = 1/n_e e$ is the Hall constant. The current component created by this electric field, that is, the "Hall current," is given by $-\mu_e \mathbf{j} \times \mathbf{B}$, where $\mu_e = \omega\tau/B$ is the electron mobility τ is the electron collision mean-free time. This leads to a more generalized statement of Ohm's Law including the Hall effect:

$$j/\sigma = (\varepsilon + \mathbf{u} \times \mathbf{B}) - \frac{\mu_e}{\sigma}\mathbf{j} \times \mathbf{B} \tag{11-40}$$

A vector diagram of the electric field components from this relation is shown in Fig. 11-69. Further theoretical discussion of the generalized Ohm's Law can be found in Refs. 9–12.

Generalized Ohm's Law. Ohm's Law, cited above, is a constitutive relationship (for instance, analogous to the equation of state for gases) and as such has a limited range of applicability. Various forms of Ohm's Law can be obtained depending on the approximations made in deriving the current-field relationships from the equations of motion and the interactions of the constituent parts of the fluid. For weakly-ionized gases in thermal equilibrium at moderate temperature, Eq. (11-40) has the equivalent tensor form (neglecting ion current):

$$\mathbf{j} = \hat{\boldsymbol{\sigma}} \cdot (\varepsilon + \mathbf{u} \times \mathbf{B}) = \hat{\boldsymbol{\sigma}} \cdot \varepsilon^* \tag{11-41}$$

or

$$j_v = \sum_k \sigma_v^k \varepsilon_k^* \tag{11-42}$$

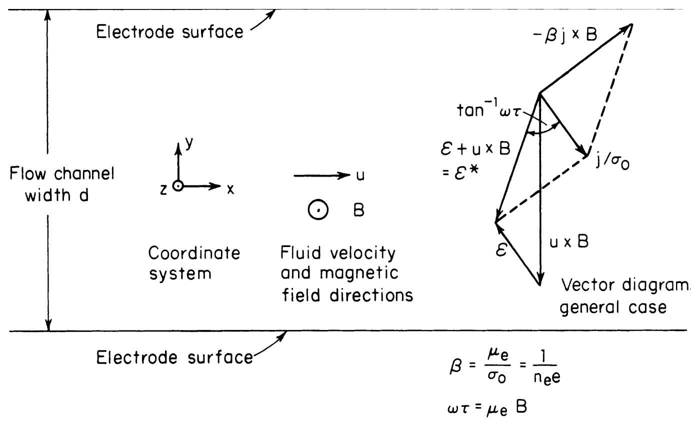


FIGURE 11-69 Vector quantities in the generalized Ohm's Law.

When $\mathbf{B} = z_1 B_z$,

$$\hat{\sigma}/\sigma f = \begin{vmatrix} \frac{1}{1 + \omega^2\tau^2} & \frac{-\omega\tau}{1 + \omega^2\tau^2} & 0 \\ \frac{\omega\tau}{1 + \omega^2\tau^2} & \frac{1}{1 + \omega^2\tau^2} & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{11-43}$$

where

$$\omega = \frac{eB}{m_e} \qquad \text{electron cyclotron frequency}$$

$$\tau = \frac{\lambda}{c_e} \qquad \text{electron collision mean free time}$$

$$\omega\tau = \mu_e B \qquad \text{Hall parameter}$$

The dimensionless product $\omega\tau$, often called the "Hall parameter," is an important characteristic number in MHD design. The conductivity tensor is anisotropic due to the Hall component unless $\omega\tau << 1$ (typical values for weakly-ionized gases are 1 to 5.) On a microscopic scale, the Hall parameter indicates the average angular travel of electrons between collisions. Typical values are $\lambda \sim 10^{-7}$ m, $c_e \sim 10^5$ m/s, $\tau \sim 10^{-12}$ s, $\omega = 1.76 \times 10^{11}\, B \sim 10^{12}$/s for $B = 6$ T. Since the mean free path is inversely proportional to pressure, lower pressure and higher values of B give larger values of $\omega\tau$.

On a macroscopic scale, the value of $\omega\tau$ indicates the relative importance of the Hall field and Hall current. When $\omega\tau = 1$, the total current is directed 45° to the left of the ε^* vector (see Fig. 11-69), and for large values of $\omega\tau$ the current vector is nearly perpendicular to ε^* (predominantly Hall current). In weakly ionized gases, if both the electron ($\omega\tau$) and ion ($\omega_i\tau_i$) Hall parameters are large simultaneously then the angle is reduced. In this case, the conductivity is reduced due to a phenomenon called "ion slip." Even though τ_i is ordinarily larger than τ_e, ω_i is much smaller than ω_e, such that the product $\omega_i\tau_i$ is usually negligible.

For highly collisional fluids (such as condensed liquids) where $\tau \to 0$, the Hall current is negligible, and the use of Eq. (11-18) without further modification is satisfactory.

Circuits with Conducting Ducts. For MHD flows in electrically conducting ducts with no external load, a return current can exist in the duct walls (Fig. 11-70). For a uniform flow velocity u, the loop voltage equation is written:

$$\oint \varepsilon \cdot dl = 2b\left(uB - \frac{j_y}{\sigma_f}\right) - 2b\frac{j_w}{\sigma_w} = 0 \tag{11-44}$$

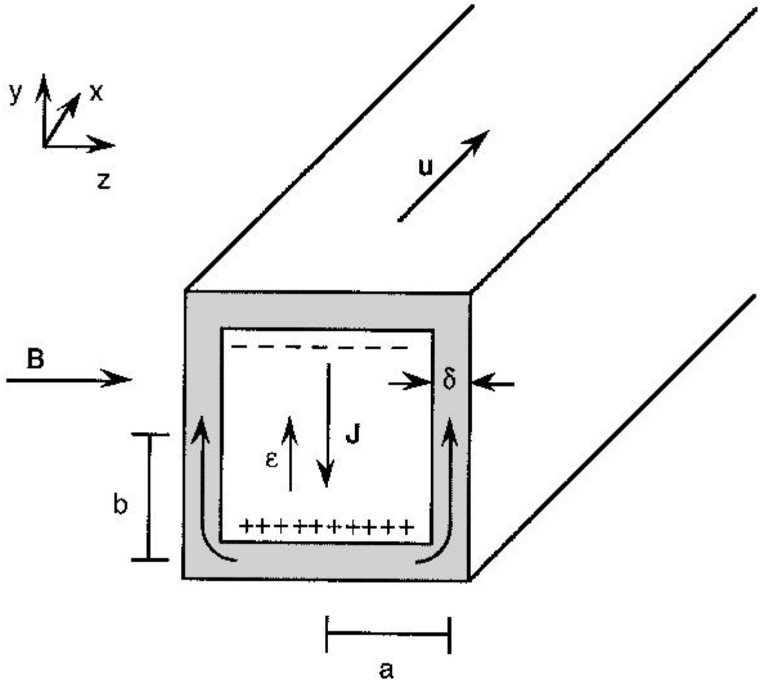


FIGURE 11-70 Current paths in a conducting duct.

where the subscript w denotes values in the wall. Conservation of current dictates:

$$j_y\, a = j_w\, \delta \tag{11-45}$$

so that Eq. (11-44) becomes

$$j_y = \sigma_f u B \frac{\Phi}{1 + \Phi} \tag{11-46}$$

where Φ is the "wall conductance ratio"

$$\Phi = \frac{\sigma_w \delta}{\sigma_f a} \tag{11-47}$$

For this type of uniform flow, Eq. (11-31) can be used to estimate the pressure required to drive a flow at velocity u through the duct.

Basic Flow Characteristics and Power Production

Hartmann Flow. Equation (11-46) predicts zero current when the walls are not electrically conducting; however, the no-slip boundary condition on the fluid at the wall, results in a nonuniform channel velocity and the formation of a boundary layer with a reduced $\mathbf{u} \times \mathbf{B}$ emf, allowing a conducting return-current path through the fluid itself.

For one-dimensional, fully developed (hence inertialess) flow, the momentum equation for $u_x(z)$ becomes

$$\frac{dp}{dx} = j_y B_z + \mu_f \frac{d^2 u_x}{dz^2} \tag{11-48}$$

where dp/dx is constant, and Ohm's Law is

$$j_y = \sigma(\varepsilon - u_x B_z) \tag{11-49}$$

It can be shown that the electric field also is constant, so that one can substitute Ohm's Law into the momentum equation and obtain a simple differential equation for u_x:

$$\mu_f \frac{d^2 u_x}{dz^2} - \sigma_f B^2 u_x = \left(\frac{dp}{dx} - \sigma_f B\varepsilon\right) \tag{11-50}$$

The solution to this equation is

$$\frac{u_x}{u_b} = \frac{\text{Ha}\cosh\text{Ha}}{\text{Ha}\cosh\text{Ha} - \sinh\text{Ha}}\left[1 - \frac{\cosh\text{Ha}\,\frac{z}{a}}{\cosh\text{Ha}}\right] \tag{11-51}$$

where u_b is the bulk average velocity. For large values of the Hartmann number, Eq. (11-51) simplifies to

$$\frac{u_x}{u_b} \approx 1 - \exp\left[\text{Ha}\left(\frac{|z|}{a} - 1\right)\right] \tag{11-52}$$

Equation (11-52) describes a velocity profile that is nearly flat throughout the duct, with thin boundary layers at the walls where viscous drag forces the flow to zero. The thickness of the Hartmann boundary layer scales as a/Ha. The shape of the Hartmann profile is shown in Fig. 11-71 for a range of Hartmann numbers.

Channel Power and Conversion Efficiency. The total electric power generated internally in a channel is equal to the mechanical work against the MHD body force, which is given by the product of the volume flow rate and pressure drop.

For a distributed medium, the Lorentz force leads to the pressure gradient

$$\nabla p = n\mathbf{F}_q = nqv \times \mathbf{B} = \mathbf{j} \times \mathbf{B} \tag{11-53}$$

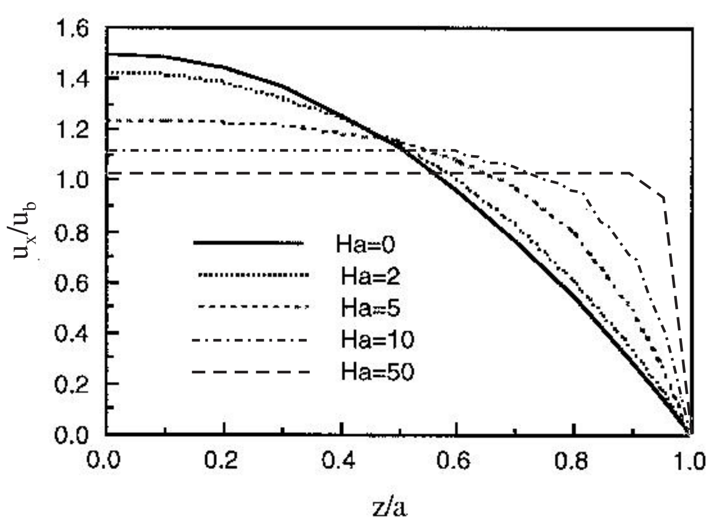


FIGURE 11-71 Normalized Hartmann velocity profiles for Ha = 0, 2, 5, 10, 50.

so that the power density is simply:

$$P_i = \mathbf{u} \cdot \nabla p = \mathbf{u} \cdot (\mathbf{j} \times \mathbf{B}) \tag{11-54}$$

The amount of power delivered to the external load is

$$P_L = \mathbf{j} \cdot \varepsilon \tag{11-55}$$

so that we can write the local electrical efficiency as;

$$\eta_e = \frac{P_i}{P_L} = \frac{j \cdot \varepsilon}{u_x j_y B_z} \tag{11-56}$$

11.10.3 Liquid MHD

Introduction. As seen above, the presence of the **j** × **B** force on the flow of conducting *liquids* can alter the velocity and pressure characteristics of the flow. The interaction with a magnetic field also can significantly affect the onset and character of turbulent fluctuations. These two effects together or individually can dramatically alter the heat-transfer characteristics and fluid drag in closed or open channel liquid flows. Technological applications of such phenomena include cooling systems for magnetic fusion reactors and reduced-drag ship hulls and airplane fuselages.

The MHD force can be applied in such a way that useful work can be done. For example, EM pumps can be designed to precisely control liquid flows, liquid metal flows, in particular, where high temperature and corrosive tendencies prohibit the use of seals in standard mechanical pumps. Such pumps have no moving parts and are extremely reliable. The converse is also possible; MHD generators can produce high currents at low voltages.

This section is concerned with exploring the interaction of the magnetic field with liquid flows both with and without applied electric currents. For incompressible liquids, the equations of Sec. 11.10.2 are valid, with Eq. (11-15) reduced to

$$\nabla \cdot \mathbf{u} = 0 \tag{11-57}$$

A discussion of various applications where such phenomena are encountered is included as well. More comprehensive discussions of many of these subjects can be found in textbooks[3,13–15] as well as the numerous other references provided throughout this section.

Closed Channel Flows

Fully Developed Channel Flow. The term *fully developed,* used to describe Hartmann flow in Sec. 11.10.2, denotes a condition where the velocity profile is no longer changing (zero derivative) in the main flow direction, that is, a flow that has reached a stable steady state driven by a constant pressure gradient, $P = -dp/dx$. The study of fully developed flow with a constant applied magnetic field is useful because the equations can be solved analytically for a variety of cases, and then can be used as benchmark problems for complete numerical algorithms. In addition, the fully developed solutions predict some phenomena of general interest, especially the existence of different boundary layers, which are important for a general understanding of MHD flows. By controlling the amount of current that can flow in the main body of the liquid, these boundary layers and the MHD boundary conditions exert a significant influence on the velocity profile and pressure drop.

Equations and Boundary Conditions for Two-Dimensional Fully Developed Flow. In a rectangular channel (a round pipe is fundamentally the same), we denote the flow direction as x and restrict the applied magnetic field B_o to be constant and aligned with z, as seen in Fig. 11-70. The MHD equations, Eqs. (11-14), (11-16) to (11-18), and (11-57), can be simplified to the following form:

$$\mu_f\left(\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + \frac{B_o}{\mu_m}\frac{\partial B_x}{\partial z} = -P \tag{11-58}$$

$$\frac{1}{\sigma_f \mu_m}\left(\frac{\partial^2 B_x}{\partial y^2} + \frac{\partial^2 B_x}{\partial z^2}\right) + B_o\frac{\partial u}{\partial z} = 0 \tag{11-59}$$

where the velocity vector has only one component in the x-direction, and the magnetic field is the sum of the constant applied field in the z-direction and the small field induced in the x-direction.

$$\mathbf{u} = [u(y, z), 0, 0] \qquad \mathbf{B} = [B(y, z), 0, B_o] \tag{11-60}$$

Terms quadratic in B have been discarded as small, an assumption equivalent to assuming small Re_m, so that B really represents a stream function for the electric current, where

$$\mu_m j_y = \frac{\partial B_x}{\partial x} \qquad \mu_m j_z = -\frac{\partial B_z}{\partial z} \tag{11-61}$$

Walls parallel to the magnetic field (i.e., "side walls") are located at $y = \pm b$, and walls perpendicular to the field (i.e., "Hartmann walls") are located at $z = \pm a$. The boundary condition for the velocity is the standard fluid-mechanical no-slip condition:

$$U = 0 \qquad \text{(at all walls)} \tag{11-62}$$

For certain boundary conditions on the induced field B, analytical solutions exist in the form of infinite series to the system. Much of the classic work in the 1950s and 1960s focused on solving these equations either exactly or approximately by expanding one dimension in an appropriate set of eigenfunctions. Some boundary conditions for the induced field are summarized in Eqs. (11-63) to (11-65). These conditions are derived by considering the behavior of the normal n and tangential s current at the wall, and using Eq. (11-61) to phrase these conditions in terms of the induced magnetic field.

$$B = 0 \qquad \text{electrically insulated wall}^{16,17} \tag{11-63}$$

$$\frac{\partial B}{\partial n} = 0 \qquad \text{perfectly conducting wall}^{18} \tag{11-64}$$

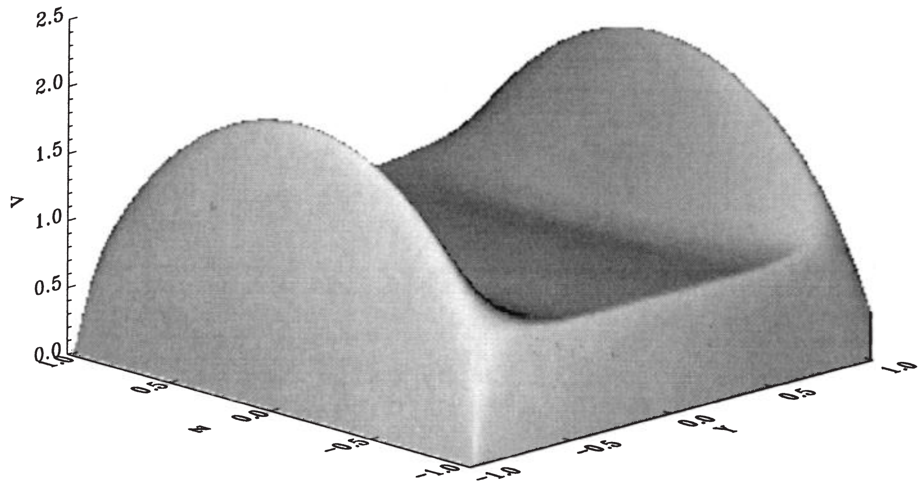
$$a\,\Phi\frac{\partial B}{\partial n} - B = 0 \qquad \text{thin conducting walls}^{19\text{–}21} \tag{11-65}$$

Inviscid Core Flow and Boundary Layers. When Ha is large, the infinite series solutions cited above show (see Fig. 11-72) the existence of a flat, inviscid core region in the central section of the duct, bordered by different types of viscous boundary layers near the Hartmann walls and side walls[22]. In this core region, the driving force of pressure is balanced entirely by the MHD force, $\nabla p = \mathbf{j} \times \mathbf{B}$. The curl of this equation implies that $(\mathbf{B} \cdot \nabla)\mathbf{j} = 0$, indicating that the current density is constant along (applied) field lines in the core. A constant pressure and a constant current produce a flat constant velocity in the core region.

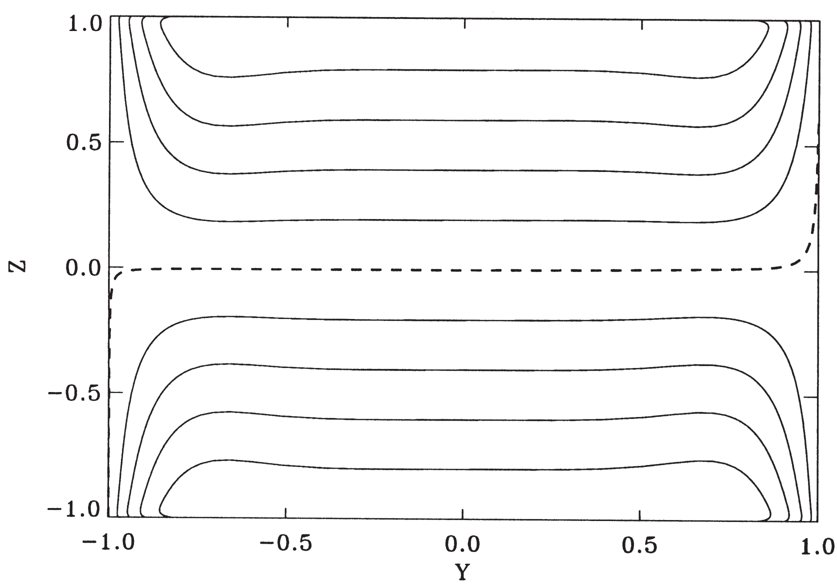
On the Hartmann walls at $z = \pm 1$, a Hartmann layer forms (similar to that seen in the one-dimensional example of "Basic Flow Characteristics and Power Production" under Sec. 11.10.2), as shown in Fig. 11-72. Hartmann layers have thickness of a/Ha, and join smoothly to the core value of the velocity. The Hartmann layer serves as a region where electric currents induced in the core flow in the y-direction can return and complete the current loop. This role as current return path makes the Hartmann layer an active boundary layer, one whose properties control the amount of flow possible in the core region. If the Hartmann wall is electrically conducting, the electric current will flow in the wall as well, and the influence of the Hartmann layer on the core flow will be accordingly reduced.

In most cases, it is the combined conductivity of the Hartmann layer and Hartmann wall that determines the MHD resistance to the fluid flow in the core, and so governs the pressure gradient P (i.e., the pressure gradient required to drive the flow at a given average velocity). In an electrically insulated channel with laminar flow, the increase in P is due to increased shear friction at the walls as a result of the modification of the parabolic laminar velocity profile. The average electromagnetic force in this case is zero since all the current induced in the flow closes through boundary layers in the fluid itself. For large Ha with insulated walls, P increases linearly with Ha. For electrically conducting channels, the net electromagnetic force is no longer zero, but can in fact be quite large. For a perfectly conducting channel with large Ha, P increases proportionally with Ha^2.

On the side walls at $y = \pm 1$, one observes the formation of a different type of boundary layer, alternately known as a "side layer," shear layer," or "Shercliff layer." The interpretation of these layers



(a)



(b)

FIGURE 11-72 Flow profiles in a square ($\beta = 1$) duct with Ha = *75*, insulated side walls, and Hartmann walls with $\Phi = 0.1$: (*a*) velocity profile; (*b*) electric current paths.

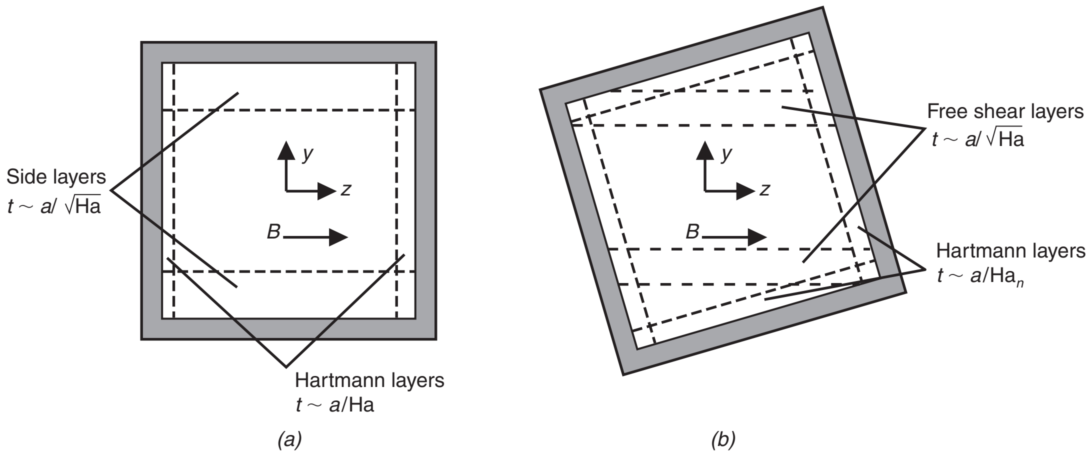


FIGURE 11-73 Boundary layers in a rectangular duct (*a*) aligned with and (*b*) oblique to an applied magnetic field.

is a region where mismatched electric potentials equalize, sometimes with a significant jet of liquid when the Hartmann walls are electrically conducting. (This is the case pictured in Fig. 11-73). These jets can carry an appreciable portion of the net mass flux under certain cases. The thickness of side layers is of order $a/\mathrm{Ha}^{1/2}$, which is much greater than that of the Hartmann layer. Thus, in most cases (except when the Hartmann walls are highly conducting but the sidewalls are not), the electrical resistance of the side layers does not add significantly to the total resistance of the return current path, and so does not influence the core velocity. For larger Hartmann numbers on the order of 10^3 or 10^4, the flow in the core region drops almost to zero, and the velocity jets can be up to a factor of Ha times the core flow velocity. Obviously, such velocity structures can be very important in determining heat transfer in liquid metal coolant pipes, as well as affecting corrosion, mass diffusion, and other important physical processes.

Hartmann layers will form on any wall that has a normal component of B_o, but shear layers form only along the magnetic field lines. Thus, if the channel is not perfectly aligned with B_o, then all walls will have Hartmann layers, and the shear layers will detach from the wall and form about the magnetic field line that intersects the corner of the duct (see Fig. 11-73*b*). Shear layers that extend into the fluid are known as *free shear layers*.[23]

Developing Flows, Variable Fields, Variable Duct Sizes, and Entrance Effects. Few practical MHD flows are fully developed over their entire length. Developing flows are inherently three-dimensional since the motion, electric currents, and magnetic field and its gradients invariably are oriented in different directions. Sudden expansion and other change in the magnetic field, channel shape, or channel electrical conductivity can result in significant changes in velocity profiles and additional three-dimensional pressure drops. Even local regions of reversed flow are possible in different MHD duct flow configurations, as in the case of a locally conducting crack in a pipe wall covered with an electrical insulator coating.

Only when the flow has advanced sufficiently far from these disturbances will it again become fully developed and assume the characteristic velocity profiles and pressure drops discussed above. A rigorous treatment of the effects of changes in field and geometry possible in MHD machines are available in the literature.[3,13,14] One example of practical interest is the entrance of a rectangular duct into a magnetic field (typical of MHD conduction pumps discussed later), and the so-called M-shaped velocity profile. Use of the core flow approximation, described in more detail below, is a powerful tool for analyzing these MHD phenomena at large Ha.

Duct Flows in Varying Magnetic Fields. Consider a rectangular duct with the orientation shown in Fig. 11-70, but assuming that at $x = 0$, the magnetic field changes abruptly from zero to some

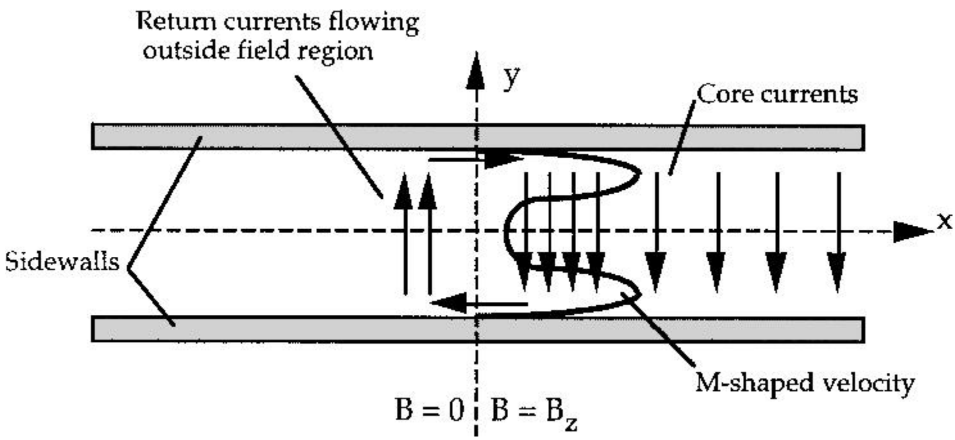


FIGURE 11-74 Simplified current paths and velocity profile for duct in a variable magnetic field.

value B_z. (The case of a more gradual transition does not fundamentally change the description.) As in the fully developed case, the $\mathbf{u} \times \mathbf{B}$ emf induces a current in the negative *y*-direction. At the sidewalls, the current can turn into the *z*-direction and then flow to the Hartmann walls, but now it can also turn (more easily) into the *x*-direction and return to the other sidewall through the region of no field directly adjacent to the region of high field (see Fig. 11-74). The same electric field set up by the charge separation that drives return current through the Hartmann layers (walls) will drive this current in the no-field region. Thus, an additional current closure path exists, and so more core current can flow in the high field region as compared to the corresponding fully developed flow in a region of constant B_z. This additional current results in a greater pressure drop, usually denoted Δp_{3D}.

The current in the *x*-direction near the sidewalls also causes a distortion of the velocity profile near the region of changing B_z. The j_x current, which is positive in the upper part of the channel ($y > 0$), will induce a force in the negative *y*-direction near the sidewall. This force will essentially pressurize the sidelayer, and cause a velocity jet to form in the sidelayer. A similar result occurs in the lower half ($y < 0$) plane. The result is a velocity profile called "M-shaped," where the velocity is reduced in the core due to increased $\mathbf{j} \times \mathbf{B}$ forces, but increased in the sidelayers. This looks very similar to the sidelayer jets that can occur in fully-developed flow when the Hartmann walls are electrically conducting, but the M-shaped profile forms in the developing region even when the entire channel is nonconducting. Like the fully developed sidelayer jets, the M-shaped velocity profile is shorted out when the sidewalls are highly conducting, since the j_x current preferentially flows in the walls in this case, and no force is induced in the liquid itself. The same effect occurs at the exit of a magnetic field, and even if the field is more gradually varied.

Mathematically, the formation of the M-shaped velocity profile can be understood by taking the curl of the steady invicid equation of motion:

$$\nabla \times \{\rho(\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \mathbf{j} \times \mathbf{B}\} \tag{11-66}$$

and considering the *z* component of vorticity:

$$\rho u \frac{\partial \omega_z}{\partial x} \cong -j_x \frac{\partial B_z}{\partial x} - B_z \frac{\partial j_y}{\partial y} \tag{11-67}$$

Both terms on the right hand side of Eq. (11-67) will be negative in the upper right quadrant of Fig. 11-74, and positive in the lower right quadrant. These sources of *z*-directed vorticity can be thought of as swirling motion that decelerates the center and shifts fluid to the sidelayers, causing the formation of the velocity jets. It is easily seen that the formation of sidelayer jets in fully developed flow is also governed by the last term in the above equation, which is present in regions of constant B_z as well.

General Core Flow Equations. For high conductivity liquids like liquid metals, calculations at high Ha can become very difficult unless simplifying approximations are made. One such approximation indicated by the above discusion is the so-called core flow approximation, where the momentum equation is simplified to

$$\nabla p = \mathbf{j} \times \mathbf{B} \tag{11-68}$$

In 1968, Kulikovskii[24] showed that the core equations (Eqs. 11-18, 11-20, 11-57, and 11-73) could be manipulated in such a way as to reduce the solution for any flow geometry without sidewalls or internal shear layers to, at most, 4 two-dimensional partial differential equations. In addition, the magnetic field is assumed to be externally applied ($\mathrm{Re}_m = 0$). The unique features of these equations allow us to separate components of velocity and current into components parallel and perpendicular to the magnetic field:

$$\mathbf{j}_\perp = \frac{B}{B^2} \times \nabla p \tag{11-69}$$

$$j_\parallel = \int \left(\mathbf{B} \times \nabla \frac{1}{B^2}\right) \cdot \nabla p \, dl + A_1 \tag{11-70}$$

$$\mathbf{u}_\perp = -\frac{1}{\sigma B^2} \nabla p + \frac{\mathbf{B}}{B^2} \times \nabla \varphi \tag{11-71}$$

$$u_\parallel = \int \left[\left(\mathbf{B} \times \nabla \frac{1}{B^2}\right) \cdot \nabla \varphi + \left(\nabla p \cdot \nabla \frac{1}{\sigma B^2}\right) + \frac{\nabla^2 p}{\sigma B^2}\right] dl + A_2 \tag{11-72}$$

$\mathbf{j}_\perp$ is obtained by taking the cross-product of $\mathbf{B}$ with the momentum equation (Eq. 11-68). Similarly, $\mathbf{u}_\perp$ is obtained by taking the cross-product of $\mathbf{B}$ with Ohm's Law [Eq. 11-18]. $j_\parallel$ and $u_\parallel$ are obtained by integrating the conservation laws (Eqs. 11-20 and 11-57 along the magnetic field direction, dl, defined by $\nabla_\parallel = d/dl$.

Finally, the electric potential can be related to the parallel current:

$$\sigma \frac{\partial \varphi}{\partial l} = j_\parallel \tag{11-73}$$

$$\sigma \varphi = \iint \left(\mathbf{B} \times \nabla \frac{1}{B^2}\right) \cdot \nabla p \, dl' dl + A_1 l + A_3 \tag{11-74}$$

The boundary conditions at the walls completely determine the unknown functions (A_1, A_2, p, and ϕ). There are four boundary conditions. Zero mass flux into the walls and conservation of current are applied twice for each field line: once where the field line enters the fluid and once where it exits.

$$\mathbf{u} \cdot \hat{n} = 0 \tag{11-75}$$

$$\Phi \nabla^2 \varphi = \mathbf{j} \cdot \hat{n} \quad \text{for conducting ducts} \tag{11-76}$$

$$\frac{1}{Ha} (\nabla \times u) \cdot \hat{n} = \mathbf{j} \cdot \hat{n} \quad \text{for nonconducting ducts} \tag{11-77}$$

This method has been successfully applied to the solution of a number of basic geometries in Ref. 25. and formulated in a very general fashion in Ref. 26. For symmetric problems, the constants A_1 and A_2 can be eliminated. The constant A_3 can be replaced by the evaluation of φ at any location along l. (For example, A_3 can be replaced by φ_w, the potential at one wall.) In this case, only two partial differential equations remain for p and φ_w.

The resulting set of linear partial differential equations can be solved using any appropriate numerical technique. A finite difference representation was applied and SOR was used to solve the resulting system of algebraic equations.[25] Corrections for sidelayers and internal shear layers also have been developed.[27]

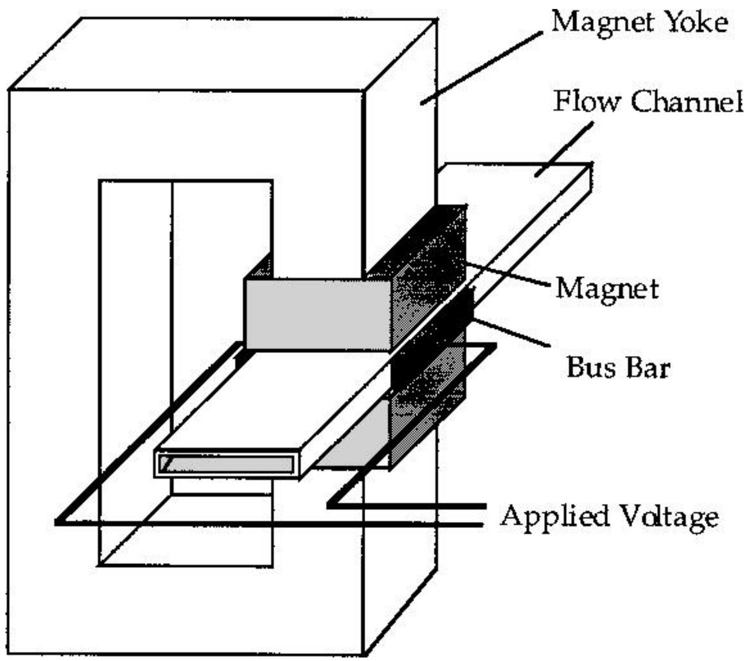


FIGURE 11-75 Generic dc conduction pump with rectangular channel.

EM Pumps and Flowmeters. One of the more practical uses of the MHD force is in pumping systems, where electrical energy is converted directly into force on the working liquid. EM pumps (as they are commonly known) have been in existence for many years, and many different designs have been successfully developed and employed. A generic conduction style pump is shown in Fig. 11-75. Another common MHD device is the EM flowmeter, where the potential induced by fluid motion is measured and used to infer the average flow rate.

These devices can be constructed with no moving parts and no direct contact with the working liquid. This is a distinct advantage if high temperature and/or corrosive liquids must be handled. The absence of seals or moving parts leads to a highly reliable system. In addition, EM pumps are typically controllable, and even reversible, by varying the magnitude and direction of the applied current.

MHD Flowmeters. Equation (11-36) suggests that the voltage induced by the $u \times B$ emf would provide an ideal method by which one could measure the average flowrate of a conducting liquid. However, it is impossible to achieve a completely open-circuit configuration as described by Eq. (11-36), as return current will flow in pipe walls and boundary layers of all finite size channels. Some accounting for these return currents must be included when determining a relation for the voltage signal as a function of the flow velocity. Ohm's Law, as written in Eq. (11-37), shows the effect of return currents on the measured voltage signal in a rectangular channel like that in Fig. 11-69:

$$-\varepsilon = \varphi/2b = uB - j/\sigma \tag{11-78}$$

where φ is the voltage signal. Any core current allowed to flow, then, reduces the measured voltage signal by some amount. Using the result of Eq. (11-46) for the core current density, we see that the voltage signal is

$$\varphi = \frac{2buB}{1 + \Phi} = \frac{QB}{2a(1 + \Phi)} \tag{11-79}$$

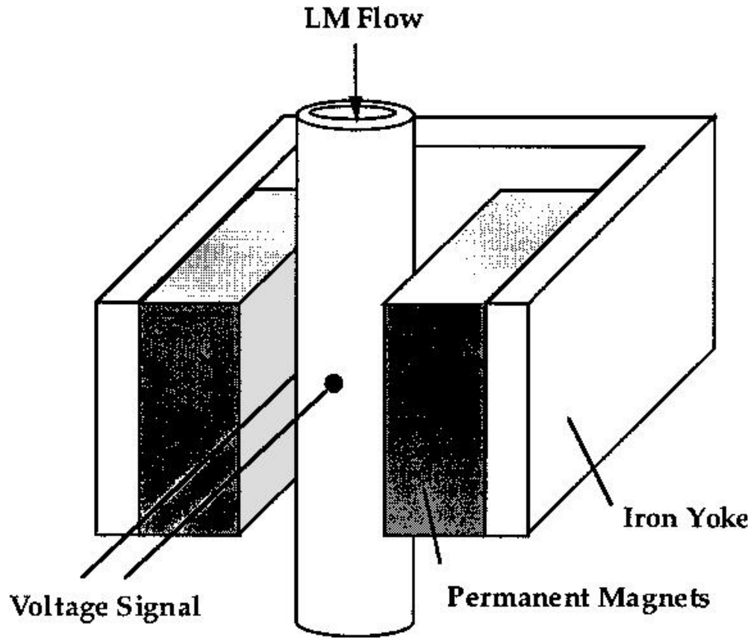


FIGURE 11-76 Generic dc EM flowmeter.

which is independent of the electrical conductivity of the liquid, except as it appears in the wall conductance ratio Φ. This means that many moderately conducting liquids can also be measured by this method, especially when an electrically insulated channel can be used. (For insulated channels, substitute Ha^{-1} for Φ in Eq. 11-79.) For common laboratory and industrial values of the flow rate Q, magnetic field induction, and duct dimensions, the measured voltage is typically in the range of hundreds of microvolts to several millivolts.

Most MHD flowmeters do not use rectangular channels, but instead use round pipes that can fit easily into standard piping systems, like that shown in Fig. 11-76. Reference 28 suggests the following relation for the volumetric flowrate in a pipe made of an electrically conducting material:

$$Q = 3162 \frac{k_4}{k_1 k_2 k_3} \frac{d\varphi}{B} \tag{11-80}$$

where Q is in units of gpm, B is in Gauss, the inside pipe diameter d is in inches, and the electric leads to measure φ are attached to the outside of the pipe wall.

Equation (11-80) takes into account, in the form of semiempirical multiplicative constants k_{1-4}, several nonideal effects that were ignored in Eq. (11-78). k_1 is the pipe-wall current shunting correction factor defined as

$$k_1 = \frac{2dD}{D^2 + d^2 + \dfrac{\sigma_w}{\sigma_f}(D^2 - d^2)} \tag{11-81}$$

where D is the outside pipe diameter, also in inches.

Poorly conducting liquids tend to have a k_1 correction factor that deviates significantly from unity, meaning that the electrical signal is lower for the same flowrate and so more difficult to measure accurately. k_2 is the magnetic field end-effect correction factor. It is empirically obtained as a function of the magnet pole piece length L. Typical values are given in Table 11-20. k_3 is the magnetic material temperature correction factor, which accounts for changes in the magnetic field as a function of temperature of the permanent magnetic material or electromagnet windings. The manufacturer of such materials typically provides the appropriate temperature correction factors. k_4 is the pipe thermal expansion correction factor, which accounts for changing pipe sizes as a function of temperature and can be expressed as $k_4 = 1 + \gamma\,(T - T_o)$, where γ is the thermal expansion coefficient for the pipe material. For 304 stainless steel, the value of γ is 9.6 μ-in/in F.

TABLE 11-20 Magnetic Field End Effect Correction Factors[28]

L/D	k_2
1.5	0.91
1.9	0.96
2.4	0.98
3.0	0.99

The idea behind these simple dc flowmeters has evolved into more complicated implementations where pulsed dc or pulsed ac electromagnets are used to sample the flowrate at some pulse rate. The pulsed electromagnet devices have lower power consumption and lower heat generation in the electric coils. These devices are available commercially[29] with all manner of pipe materials and sizes and can be inserted easily into existing piping configurations. Small units ($d \approx 1$ in) have accuracies of around 1% at full scale. The accuracy tends to improve as pipe sizes become larger.

Conduction Pumps. Liquid metal conduction pumps, or "Faraday" pumps, consist of a rectangular channel in the gap of a magnet (either a permanent or electromagnet) where an electric current is passed through the conducting liquid perpendicular to the field. The resulting $\mathbf{j} \times \mathbf{B}$ force drives the flow. This type of pump is a conceptually simple extension of the rectangular duct flows discussed above with sidewalls replaced by electric bus bars connected to an outside voltage source, which drives current in the direction opposite to the $\mathbf{v} \times \mathbf{B}$ emf.

To understand the pressure and flowrate behavior of a simplified conduction pump, pictured in Fig. 11-75, it is helpful to represent the pump by an equivalent electrical network Fig. 11-77. Here, R_{LM} is the effective resistance of the liquid in the channel, R_{Loss} is the resistance of any loss paths such as the channel walls and fringing paths outside the magnetic field area, and E_i is the voltage induced in the flow, which works against the applied current I_a. Assuming slug flow of flow rate Q in a rectangular pumping channel of flow length L within the field, with walls of thickness δ_w and electrical conductivity σ_w, the following approximations can be applied:

$$R_{LM} = \frac{b}{La\sigma_f} \qquad R_{\mathrm{Loss}} = \frac{b}{L\delta_w\sigma_w} \tag{11-82}$$

$$I_{LM} = \frac{2a\Delta p}{B} \qquad E_i = 2buB = \frac{QB}{2a} \tag{11-83}$$

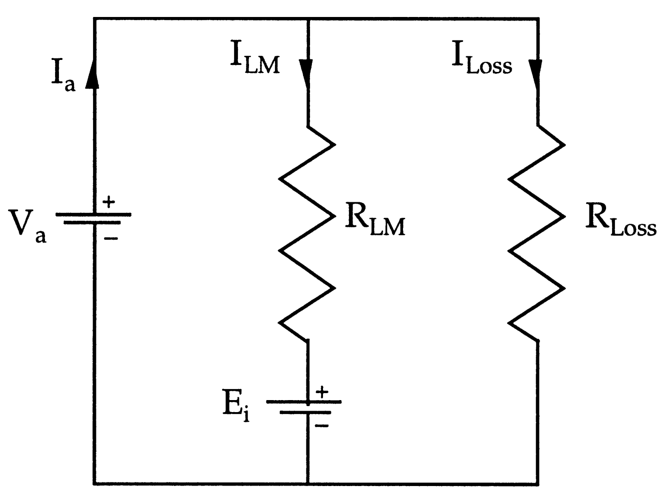


FIGURE 11-77 Circuit network equivalent of a conduction pump.

The duct geometry used is the same as in Fig. 11-70, and the R_{loss} term only takes into account current losses through conducting Hartmann walls for simplicity. The circuit then can be solved to give the following linear relationship between pressure and flow rate:

$$\Delta p = \frac{BAI_a/2a - B^2LQ\Phi\sigma_f}{A(1 + \Phi)} \tag{11-84}$$

where $A = 4ab$ is the flow cross-sectional area, and Φ is the wall conductance ratio defined in the usual way as in Eq. (11-47). This relationship is plotted in Fig. 11-78 for a small conduction pump with various values of Φ. We see the obvious need for thin and/or low conductivity walls in the pumping channel structure in order to maximize the pumping power. In the case of $\Phi = 0.1$, at flowrates above 7.51/s, all of the applied current is simply shunted through the walls, as the induced emf is larger than the applied voltage. Current flow in the core is reversed, and MHD drag, instead of pumping, results.

The applied voltage for this particular pump, assuming $Q = 10$ l/s and $\Phi = 0.01$, is only 213 mV. The inherently low voltage and high current of conduction pumps is one of their disadvantages, since they require special power supplies capable of coupling efficiently to such a load. Possible power supplies that are more efficient than standard transformer-rectifier systems include homopolar and unipolar generators, which can be up to 80% efficient. To increase the voltage of the load somewhat, it is common to run the field coil windings of the electromagnet in series with the LM section, so that only one supply (albeit at a greater power level) is needed.

The power dissipated in the conduction pump system for our example above is

$$P_a = 0.215 \text{ V} \cdot 1500 \text{ A} = 322 \text{ W} \tag{11-85}$$

The power delivered to the fluid ($Q \cdot \Delta p$) is $P_f = 0.01 \text{ m}^3/\text{s} \cdot 25.7 \text{ kPa} = 257$ W. This gives an efficiency of 81%. A formula for the electrical efficiency in the general case is easily constructed in terms of either the applied voltage or the applied current:

$$\eta_e = \frac{Q\Delta p}{V_a I_a} = \frac{BQ}{2aV_a\left(1 + \dfrac{\Phi}{1 - BQ/2aV_a}\right)} = \frac{2bI_a - BLQ\sigma_f\Phi}{2bI_a\left(1 + \dfrac{2bI_a}{BLQ\sigma_f}\right)} \tag{11-86}$$

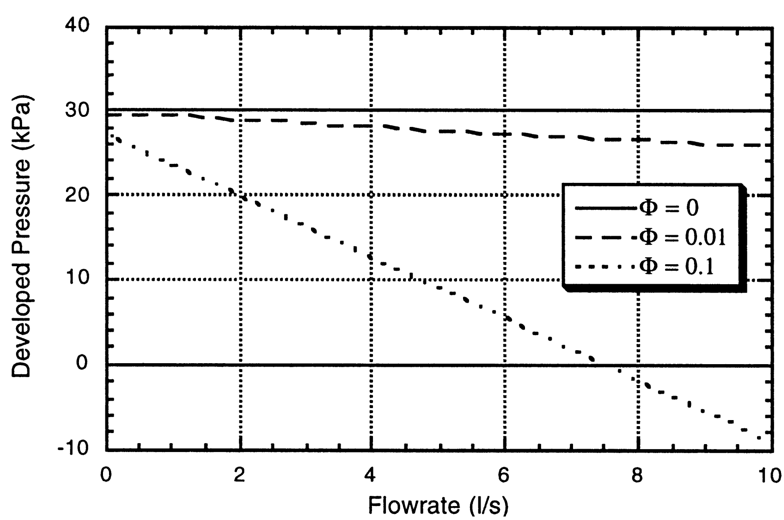


FIGURE 11-78 Pressure head of a conduction pump for various wall conductance ratios and the following parameters: $a = 5$ cm, $b = 15$ cm, $L = 30$ cm, $B = 2\ T$ (iron core electromagnet), $I_a = 1{,}500$ A, and $\sigma_f = 10^6\ \Omega$ m.

However, the losses in the wall included in this simple calculation are not the only losses the conduction pump experiences. Applied current usually fringes outside the area of the applied magnetic field where it induces no $\boldsymbol{j} \times \boldsymbol{B}$ force. Energy is lost to the damping of turbulence and restructuring of the average velocity profile as the liquid enters the magnetic field area. Energy losses occur due to friction of the flow on the channel walls and the applied magnetic field is altered by the field induced by the applied current itself. All of these effects decrease the net efficiency of conduction pumps to the order of 15% to 20% for small pumps, and 40% to 50% for larger pumps. Including losses in generators and field windings, the bus bar efficiency for liquid metal conduction pumps varies from 10% to 40%, increasing with pump size.[30]

As seen from Eq. (11-86), even in the absence of losses other than the resistance of the LM itself, the electrical efficiency of the system is equal to $\eta_e = \eta_{\text{ind}} = 2buB/V_a$, which is the induction efficiency, that is, the ratio of induced voltage due to the $\boldsymbol{u} \times \mathbf{B}$ emf to the applied voltage V_a.

Induction Pumps. An alternative to the conduction pump is the induction pump, where electric currents are induced in the liquid metal by means of a time-varying magnetic field, producing a $\boldsymbol{j} \times \boldsymbol{B}$ force with the instantaneous field to drive the flow. Many types of induction pumps are possible. Here we focus on the flat linear induction pump and the annular linear induction pump. The advantage of induction pumps is that they can be driven easily by single-phase or 3-phase ac power sources, possibly with a variable transformer for control of the flow rate. Typical disadvantages are greater power losses and the need for electrical insulation at high temperatures close to the working liquid.

The flat linear induction pump, or FLIP is conceptually similar to an ac induction motor. The 3-ϕ winding, however, produces a sliding, rather than rotating magnetic field, which tends to pull the fluid along. The action of this class of pump is easily pictured by contemplating the simplified induction pump shown in Fig. 11-79. If the peak vertical magnetic field is sliding to the right, leaving behind a slightly reduced magnetic field, a current loop will be induced. This induced current tries to maintain the field at its original strength. The induced current into the page will be in a region of stronger magnetic field than the current coming out of the page, since the field peak is propagating to the right, so the net $\boldsymbol{j} \times \boldsymbol{B}$ force will be to the right.

The disadvantage to these systems is that in order to have a relatively wide channel for liquid flow, the gap between the stators must be larger than that in induction motors, and the field losses are

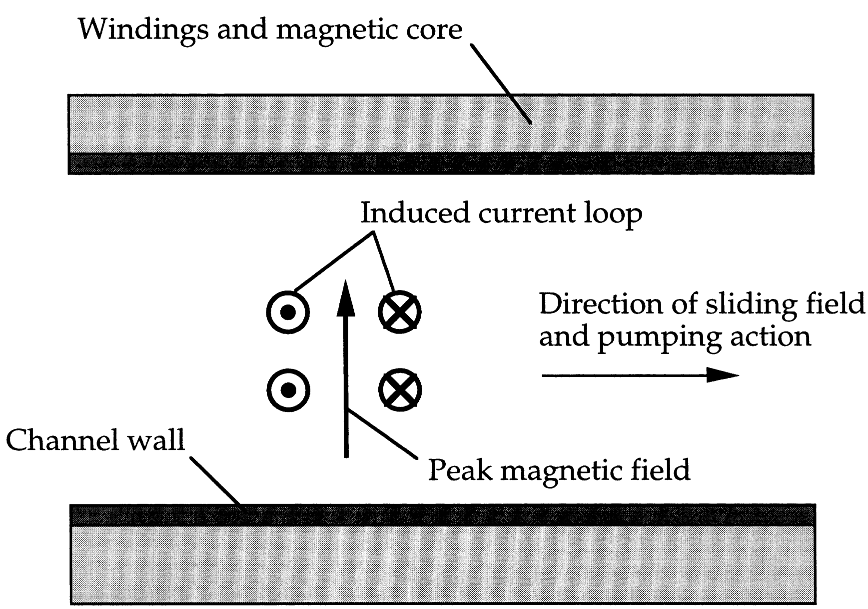


FIGURE 11-79 Operation of an induction pump.

relatively higher. In addition, field fringing occurs at the boundaries of the wide, flat channel where the magnetic core stops. One-sided stator induction pumps are also possible, but losses will be even higher. Such systems have uses as EM stirrers and for ship propulsion, as will be discussed later.

The annular linear induction pump (ALIP), sometimes called an Einstein-Szilard pump, is a modification of the FLIP so that end effects do not cause additional field losses. The ALIP consists of an annular flow region with an internal magnetic core. Induced currents flow in a continuous loop through the liquid, and no core edge exists off which magnetic fields fringe. In essence, an ALIP is like a FLIP that has been bent around until the free ends are joined.

Design guidelines and photographs of various styles of EM induction pumps are available in Ref.[31].

MHD Ship Propulsion. Another potential use of EM pumping technology is MHD thrusters for ship propulsion. Seawater has a moderate electrical conductivity, of the order of 5 $(\Omega \cdot \text{m})^{-1}$, and under the appropriate set of conditions can be pumped by the Lorentz force. Care must be taken to avoid large losses in conducting walls in this application, but this is more easily done when the working fluid is seawater rather than high temperature liquid metals.

Conduction pump thrusters[32] are more commonly envisioned for MHD ship propulsion because of the difficulty inducing large currents in poorly conducting water. Using the above equations for the conduction pump, we find that the ideal conduction pump thruster will deliver a power to the liquid equal to

$$P_w = F_{\text{EM}}u = (1 - \eta_e)\eta_e \frac{\sigma_f V_a La}{b} \tag{11-87}$$

For a given size channel (usually limited by the size of the craft under consideration), a given applied voltage (usually limited by the power supply aboard the craft, e.g., a battery), and a given liquid (seawater), the mechanical power is maximized at $\eta_e = 50\%$. This means one-half of the electrical power supplied is lost as Ohmic heating. Thruster designers must decide whether their goal is to maximize mechanical power or to minimize energy consumption.

For a moderately sized submarine (10-m diameter, 83-m length) using four conduction thrusters with length $L = 55$ m, $b = 5$ cm, and $a = 15$ cm, a 5-T field is sufficient to generate reasonable thrust and efficiency. At a speed of 36 knots, the thrusters will consume about 66 MW of electric power, requiring a 200 MW thermal nuclear plant with a typical thermal conversion efficiency (excluding power needed for other boat systems). This level of power is not unreasonable for a submarine of this size. Superconducting magnets are necessary for this field strength and core size, since the Ohmic losses in resistive magnets would be unacceptable.

At least one design using an induction pump thruster has been advanced. The "ripple motor" described by Mitchell and Gubser[33] utilizes a 3-ϕ ac solenoidal winding around a core of seawater. An annulus of liquid sodium or other liquid metal serves as an intermediate layer separated from the seawater by a flexible membrane. The thickness of the sodium layer is matched to the skin depth of the ac field. The traveling magnetic field sets up a traveling pressure wave in the sodium and thus a traveling wave on the flexible membrane. This wave pushes along the seawater and eventually ejects it out of the trailing end of the thruster, providing the thrust.

Turbulence in Liquid MHD Flow. MHD forces can have a large effect on the turbulence structure of liquid flows. Not only does the induction of a current density result in Ohmic dissipation of energy, a new energy loss mechanism that augments the viscous dissipation, but the field also changes the average velocity profiles as discussed in previous sections, resulting in new turbulence-creation scenarios as compared to non-MHD flows. The magnetic field is typically thought to *laminarize* already turbulent flows or to prevent the transition to turbulence in laminar flows. In electrically conducting channels,[13] core velocity fluctuations are damped for values of Ha/Re $>$ 0.008. Near the sidelayer jets, though, turbulent fluctuations increase, indicating that the strong velocity jets, like those depicted in Fig. 11-71, are unstable and periodically break down. For Ha/Re $>$ 0.02, these fluctuations are also damped (or at least unresolved due to boundary layer thinning), and the liquid flow becomes effectively laminar.

Electrically insulating channels exhibit a change, as the field is applied, from standard turbulence to a quasi-two-dimensional turbulence, where the vorticity of the turbulent fluctuations is predominantly aligned with the direction of the field. Turbulence fluctuations of this type can be quite long-lived and probably result from the reorganization of the flow as it enters into a magnetic field. The formation of M-shaped velocity structures, which then decay into rotating vortices, is the source of such fluctuations. In an infinitely long, electrically insulated channel, all turbulence fluctuations are eventually damped when Ha/Re > 0.008.

Control of turbulence near the wall of a ship or submarine can in theory reduce the overall drag on the structure. Early work on MHD channel flows[34] showed that the pressure drop in an initially turbulent LM duct flow could be reduced by the judicious application of a magnetic field. (Too strong a field will result in increased MHD drag, as discussed above.) For the control of turbulence near ships, one must contend with the fact that seawater is a poor electrical conductor, and that induced currents alone will not dissipate enough energy to stabilize a turbulent boundary layer. Instead, currents must be generated by an applied voltage.

One such scheme to reduce drag on, and radiated noise from, a flat plate is to construct the surface with alternating north and south pole magnets interspersed between positive and negative electrodes (see Fig. 11-80). The crossing lines of magnetic field and current induce a $\boldsymbol{j} \times \boldsymbol{B}$ force in the streamwise direction, acting as a sort of one-sided conduction pump. Preliminary experiments[35] have shown that turbulent fluctuations can be reduced over much of the boundary layer when the modified interaction parameter ($N^* = j_o B_o \theta / 0.5 \rho u_\tau^2$, where j_o, B_o are the current, field at the electrode, magnet surface, and θ and u_τ are the standard momentum thickness and friction velocity of the boundary layer, respectively) is order one or larger. The boundary layer is found to approach an asymptotic value, rather than growing indefinitely and breaking down due to instability. Work in this area by a number of researchers is continuing.

Open Channel Flows. Open channel flows of liquids in magnetic fields are of interest for metallurgical and welding applications where melts and melt layers are influenced by electric currents and applied magnetic fields. There is also interest in open channel MHD flows in magnetic fusion energy reactors where it might be advantageous to have high heat flux surfaces facing the burning plasma be covered with a flowing liquid metal layer.[36] When the problem of open channel MHD flows is examined closely, one finds that the complicated motion of closed-channel duct flows described above becomes even more complicated when the liquid interface (free surface) is allowed to move in response to MHD forces.

The interfacial boundary condition for open-channel flows requires that the tangential component of the viscous stress must be continuous. The term "free surface" implies the less general case where the liquid surface is unhampered by friction with a gas phase outside the liquid region, and so the tangential component of the stress vanishes. However, this condition is changed in MHD flows

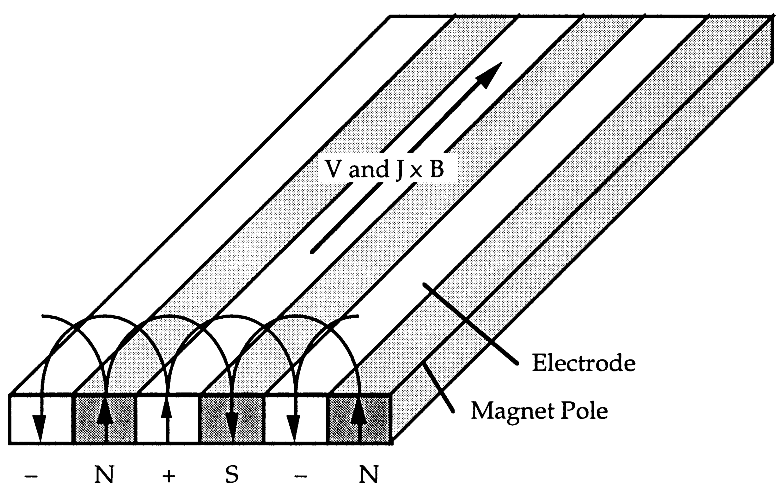


FIGURE 11-80 Turbulence reduction technique for seawater: alternating magnets and electrodes.

where the total stress, the sum of the tangential viscous and magnetic stresses, must be continuous. The magnetic stress is represented by the Maxwell stress tensor (found elsewhere in this handbook), and can exist in vacuum as well as in conducting media. This EM stress vanishes for temporally and spatially uniform magnetic fields, but needs to be considered in the general case.

Some simple cases of flow down an inclined plate are analyzed by Alpher,[37] Aitov,[38] Morley[39] and others. It is seen that for a magnetic field normal to the surface of a very wide, long plate, a half-Hartmann velocity profile forms in the surface normal direction. The flow is essentially that shown in Fig. 11-70 split open, where $z = 0$ is the free surface and $z = 1$ is the back plate. The Hartmann layer on the plate is exactly the same as one would expect in channel flow, and provides a return-current path for currents generated in the core. Also similar to flow in ducts, the modification of the velocity profile and the unbalanced $\boldsymbol{j} \times \boldsymbol{B}$ force can cause an increase in drag. For flow down an inclined plate, where there can be no applied pressure driving the flow, this results in a thickening and slowing down of the flow. More complicated magnetic field and flow geometries are considered in Ref. 40 where the effects of magnetic field gradients are especially emphasized.

Similar to the effect of magnetic fields on turbulence is the stabilizing effect of magnetic fields at the free surface. It has been shown both theoretically and experimentally that a constant strong magnetic field can stabilize an otherwise wavy free surface, resulting in a smooth flow. For the flow described in the preceding paragraph, Hsieh[41] found that for high Ha, the surface is stable to long wavelengths provided that

$$\text{Re} < \frac{\exp(2\text{Ha})}{4} \cot\theta \qquad (11\text{-}88)$$

where Re is the Reynolds number of the flow and θ is the angle of inclination of the plate to gravity. This is a much greater range of Reynolds number than the classical non-MHD result of Re $<$ 5/4 cot θ.

Applications in Metals Processing. Metals processing requires the handling of large amount of liquified metals in a controlled manner. Certainly the MHD devices discussed above, for example, pumps and flowmeters, will have manifold applications in this industry. But it is also possible to actively control the shape of a free surface by use of high-frequency ac magnetic fields.

A high-frequency magnetic field in the region around an electrical conductor, like a liquid metal, takes a finite time to penetrate into the conductor. It is easily seen from Eq. (11-23) in the limit of slow motion of the liquid $\boldsymbol{u}$ as compared to the sinusoidal oscillation frequency ω, that:

$$\frac{B}{\tau} \cong \frac{1}{\sigma_f \mu_m} \frac{B}{\delta^2} \qquad (11\text{-}89)$$

If the characteristic time τ is taken as $2\omega^{-1}$, then the skin depth $\delta \cong (2/\sigma_f \mu_m \omega)^{1/2}$. This means that if δ is small, the field cannot penetrate far into the conducting medium during the oscillation period of the ac field. In reality, currents induced in the skin region act to nullify the applied field variation. The resulting $\boldsymbol{j} \times \boldsymbol{B}$ force can have both a pressure-like component and a tangential stress component. The pressure component can be applied to the free surface in such a way to deflect and shape jets of liquids issuing from a nozzle and even levitate an entire melt. The tangential stress components can be used to induce motion and stir the melt as desired. The induced currents can also provide significant Joule heating in the skin region.

Using a combination of all these effects, it is possible to design various MHD devices like levitating, self-stirred, induction furnaces where the LM never comes in contact with a solid crucible surfaces, and MHD granulators where free LM jets or sheets decay into droplets that then solidify into powders. The interested reader is referred to Davidson,[3] Moreau,[13] and Kolesnichenko[2] for more detailed mathematical descriptions of these problems with more complete bibliographies.

11.10.4 Gaseous MHD

Introduction. Since 1959, substantial effort has been devoted to exploring the conditions under which a conducting gas moving through a magnetic field might generate useful electrical power. The primary motivation for the development and use of MHD generators in central-station power plants is the production of power at lower cost through reduced fuel costs per unit of energy produced, traded off against additional capital and operating costs. Operation at high thermal conversion efficiency provides the added benefit of reduced thermal discharge from the plant, thus reducing thermal pollution as compared with conventional steam plants with $\eta \sim 40\%$ or nuclear plants with efficiency as low as 30%.

As originally envisioned, the MHD generator was a "topping" unit on an otherwise conventional steam turbine-generator station. In this case, electric power is generated in the MHD unit, and its exhaust heat, with temperature as high as 2,200 K, is used to generate steam. The limiting Carnot efficiency for such a station might be raised from a maximum of about 65% ($T_1 = 850$ K, $T_2 = 300$ K) upward toward 85% ($T_1 = 2{,}600$ K, $T_2 = 420$ K). The net efficiency of the combined cycle can be expressed as $\eta_1 + \eta_2(1 - \eta_1)$, where η_1 is the efficiency of the MHD generator and η_2 is the efficiency of the "bottom" steam plant. Typical values are $\eta_1 = 0.25$ and $\eta_2 = 0.4$, for an overall efficiency of 0.55.

Perhaps the greatest importance of the MHD steam plant, as now envisioned, is its potential for very low air pollution while burning high-sulfur coal.[42] The SO_2, NO_2, and particulate emissions are all reduced to very low levels by their interaction with the MHD "seed" material. In pilot plant tests, 2.2 wt.% sulfur coal was burned in a cyclone furnace at 2,200°C with seed concentration of 1 g · mole K_2CO_3/kg coal with 99.8% removal of SO_2, leaving only 5 ppm SO_2 in the gaseous effluent. This occurs because of an affinity that the potassium seed material has for SO_2. So seed recovery in the MHD system, which is necessary for economic reasons, also removes the SO_2. The seed removal costs are calculated as approximately one-fifth of the SO_2 removal costs in a conventional coal-fired plant. In the same tests, through the use of 2-stage combustion, NO_x emissions were reduced below 150 ppm, and complete combustion of CO was achieved.[42,43]

Table 11-21 summarizes the main features of both open and closed cycle systems which remain the subject of both theoretical and experimental investigation. Generally, practical designs have dc output taken from electrodes at the sides of the MHD channel. Ac power is obtained using electronic inverters.

In the remainder of this section, we first review the main generator configurations used in MHD power generation, together with their governing electrical equations. Flow behavior is described for an example case, using the segmented-electrode Faraday type of generator. Following this, properties of seeded gases are given, and finally the major engineering issues for electrodes and magnets are summarized.

Generator Configurations. Figure 11-81 depicts the basic elements of a generic MHD generator, having quasi-one-dimensional flow of partially ionized gas channeled through a perpendicular, static magnetic field. For this geometry, $\mathbf{u} = \hat{\mathbf{x}}\, u_x$ and $\mathbf{B} = \hat{\mathbf{z}} B_z$. In the general case, the components of Ohm's Law (Eq. 11-22) are

$$j_x = \sigma\varepsilon_x - \mu_e j_y B_z \tag{11-90}$$

$$j_y = \sigma\varepsilon_y - \sigma u_x B_z + \mu_e j_x B_z \tag{11-91}$$

$$j_z = \varepsilon_z = 0 \tag{11-92}$$

TABLE 11-21 Gaseous MHD Power Generating System Features

	Open cycle	Closed cycle
Heat Source	Coal Manufactured gas Natural gas H_2 Fuel oil	Gas-cooled nuclear reactor Coal Natural gas Fuel oil
Working Fluid	Potassium-seeded combustion products	Cesium-seeded helium
Temperature	~2500°C	~1400°C
Magnetic Field Source	DC superconducting magnets, 4–6 T	DC superconducting magnets, 4–6 T

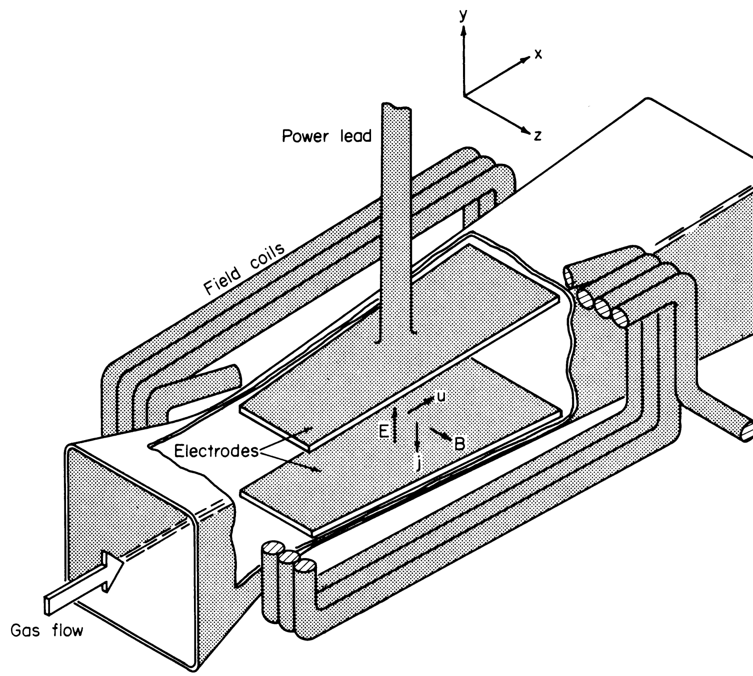


FIGURE 11-81 Basic elements of a magnetohydrodynamic generator.

Four alternative generator configurations are considered here and depicted in Fig. 11-82:

(*a*) Segmented-electrode Faraday generator
(*b*) Continuous-electrode Faraday generator
(*c*) Hall generator
(*d*) Diagonally connected generator

(a) Segmented-electrode Faraday generator. Configurations in which the circuit for j_y is completed through the external load are called *Faraday generator* configurations. If the electrodes are segmented along the *x*-direction in order to electrically isolate each pair, then the Hall current is suppressed ($j_x \approx 0$).

Assuming uniform conditions across the channel, the open circuit voltage ($j_y = 0$) is given by Eq. (11-91):

$$V_{oc} = -\int_o^d \varepsilon_y \, dy = -\int_o^d u_x B_z \, dy = -u_x B_z d \tag{11-93}$$

and the load power generated is

$$P_L = j_y \varepsilon_y = -\sigma(u_x B_z - \varepsilon_y)\varepsilon_y = -\sigma u_x^2 B_z^2 (1 - K)K \tag{11-94}$$

where the dimensionless loading parameter K is defined by $K = \varepsilon_y / u_x B_z$. Using Eq. 11-56, we find that the conversion efficiency is

$$\eta_e = \frac{j_y \varepsilon_y}{u_x j_y B_z} = \frac{\varepsilon_y}{u_x B_z} = K \tag{11-95}$$

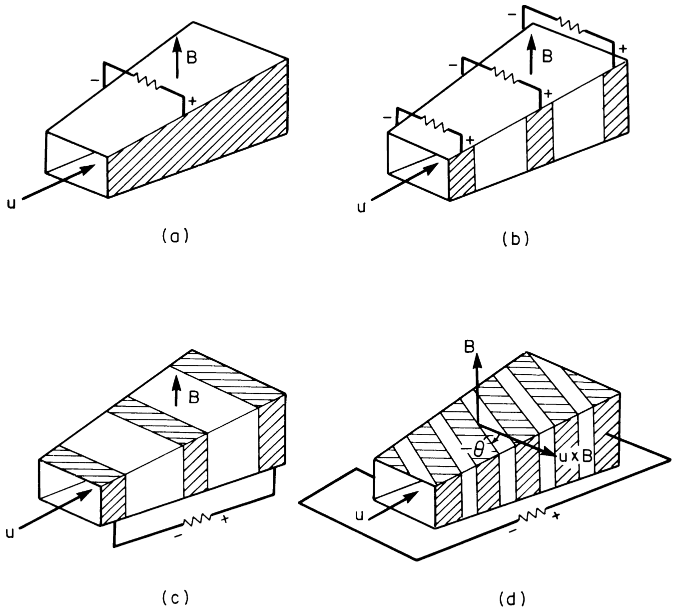


FIGURE 11-82 Generator configurations: (*a*) segmented-electrode Faraday generator; (*b*) continuous-electrode Faraday generator; (*c*) Hall generator; and (*d*) diagonally-connected generator.

(b) Continuous-electrode Faraday generator. In a continuous-electrode generator, $\varepsilon_x = 0$ and the Hall current is finite, $j_x \neq 0$. The *x*-component of current (in the direction of the fluid flow) has its circuit completed through the electrode walls. The j_y component is reduced due to this effect:

$$j_x = -\mu_e j_y B_z$$

$$j_y = \frac{\sigma}{1 + \omega^2\tau^2}(\varepsilon_y - u_x B_z) = \frac{\sigma u_x B_z}{1 + \omega^2\tau^2}(1 - K) \qquad (11\text{-}96)$$

The load power is:

$$P_L = j_y \varepsilon_y = \frac{\sigma}{1 + \omega^2\tau^2}(\varepsilon_y - u_x B_z)\varepsilon_y = -\frac{\sigma u_x^2 B_z^2}{1 + \omega^2\tau^2}(1 - K)K \qquad (11\text{-}97)$$

The open-circuit terminal voltage is the same for both types of Faraday generator, as is the local electrical efficiency, $\eta_e = K$. However, the power density is reduced by $1 + \omega^2\tau^2$ when the electrodes are continuous.

(c) Hall generator. In the Hall generator configuration, opposing electrode pairs are short circuited ($\varepsilon_y = 0$), and the circuit for j_x is completed through the external load. In this case, the current components derived from Eqs. (11-90) to (11-92) are

$$j_x = \frac{\sigma}{1 + \omega^2\tau^2}(\varepsilon_x + \omega\tau u_x B_z) \qquad (11\text{-}98)$$

$$j_y = \frac{\sigma}{1 + \omega^2\tau^2}(\omega\tau\varepsilon_x - u_x B_z) \qquad (11\text{-}99)$$

The open circuit voltage ($j_x = 0$) across the full channel length is given by

$$V_{oc} = -\int_o^L \varepsilon_x \, dy = \int_o^L \omega\tau u_x B_z \, dx = \omega\tau u_x B_z L \tag{11-100}$$

In this case, the Hall loading parameter is written $K_H = -\varepsilon_x/\omega\tau u_x B_z$, so that the load power generated is

$$P_L = j_x \varepsilon_x = \frac{\sigma}{1 + \omega^2\tau^2}(\varepsilon_x + \omega\tau u_x B_z)\ \varepsilon_x = \frac{\sigma\omega^2\tau^2 u_x^2 B_z^2}{1 + \omega^2\tau^2}(1 - K_H)K_H \tag{11-101}$$

The conversion efficiency is

$$\eta_e = \frac{j_x \varepsilon_x}{u_x j_y B_z} = \frac{(1 - K_H)K_H}{K_H + 1/\omega^2\tau^2} \tag{11-102}$$

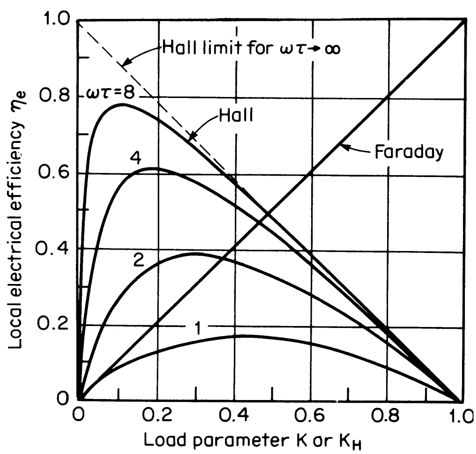


FIGURE 11-83 Comparison of local electrical efficiencies of Faraday and Hall generators.

A comparison between Faraday and Hall generator efficiencies is given in Fig. 11-83. Good efficiency in a Hall generator requires high $\omega\tau$ and a small loading parameter, whereas the Faraday generator efficiency is independent of $\omega\tau$ and improves with higher values of the loading parameter.

(d) Diagonally-connected generator. A configuration that has been favored recently is diagonal connection of electrodes along equipotential surfaces at the angle $\theta = \tan^{-1}\varepsilon_y/\varepsilon_x$ with respect to the vector $\mathbf{u} \times \mathbf{B}$. Ideally, with this configuration, the Hall current is zero, and the electrode voltage between opposite pairs is the same as for the Faraday generator. With the diagonal connections, the overall circuit is a series connection of multiple Faraday-generator electrodes. The output, at comparatively high voltage, is taken between the first and last electrodes in the series.

The characteristics of the diagonally connected generator are intermediate between those of the Hall and Faraday generators. Allowing for a finite Hall current, the current components in their most general form are

$$j_x = \frac{\sigma}{1 + \omega^2\tau^2}[\varepsilon_x - \omega\tau(\varepsilon_y - u_x B_z)] = \frac{\sigma u_x B_z}{1 + \omega^2\tau^2}\omega\tau\,(1 - K) - \frac{K}{\tan\theta} \tag{11-103}$$

$$j_y = \frac{\sigma}{1 + \omega^2\tau^2}(\omega\tau\varepsilon_x + \varepsilon_y - u_x B_z) = -\frac{\sigma u_x B_z}{1 + \omega^2\tau^2}\left(1 - K + \frac{\omega\tau K}{\tan\theta}\right) \tag{11-104}$$

The condition $j_x = 0$ is satisfied if

$$K = \frac{\omega\tau\tan\theta}{1 + \omega\tau\tan\theta} \tag{11-105}$$

In this case, the local electrical efficiency is $\eta = K$.

Energy Extraction and Flow Relations. In general, numerical solutions are needed to solve the complete set of equations governing power generation and flow. The simple case of a quasi-one-dimensional constant-velocity ideal gas flow is examined here in closed form in order to provide insight into the flow behavior. The velocity is maintained constant as the gas expands by adjusting the flow area Ac such that mass conservation leads to

$$\frac{d}{dx}(\rho A_c) = 0 \tag{11-106}$$

A Faraday generator configuration with segmented electrodes is assumed. In this case, the Hall current vanishes and Ohm's Law can be written as

$$j = \sigma(\varepsilon - uB) = -\sigma uB(1 - K) \tag{11-107}$$

The pressure variation is linear for a constant field:

$$\frac{dp}{dx} = jB = -\sigma uB^2(1 - K) \tag{11-108}$$

The solution can be written in terms of the pressure at the inlet ($x = 0$):

$$\frac{p}{p_o} = 1 - \frac{x}{L_i} \tag{11-109}$$

$$L_i = \frac{1}{1 - K}\frac{p_o}{\sigma uB^2} \tag{11-110}$$

where L_i is the "interaction length,"[44] which is an approximate measure of the channel length required to extract an appreciable amount of the gas energy.

The energy equation can be used to determine the temperature variation along the duct. The total fluid enthalphy (per unit mass) is the sum of the kinetic energy, internal energy, and static pressure:

$$H = u^2/2 + e + p/\rho \tag{11-111}$$

Conservation of energy equates the rate of change of fluid energy with the electric power dissipated:

$$\rho\frac{dH}{dt} = \mathbf{j} \cdot \varepsilon \tag{11-112}$$

With constant velocity, we can replace H by the static enthalpy h:

$$h = H - u^2/2 \tag{11-113}$$

For an ideal gas, $h = C_pT$, so that the energy equation becomes:

$$\rho C_p u\frac{dT}{dx} = j\epsilon \tag{11-114}$$

We can replace the current density and electric field in Eq. (11-114) using Eq. (11-31) and the definition of the loading parameter, $K = \varepsilon/uB$:

$$\rho C_p u\frac{dT}{dx} = \frac{1}{B}\frac{dp}{dx} \cdot uBK \tag{11-115}$$

Solving for K,

$$K = \rho C_p\frac{dT}{dp} \tag{11-116}$$

For an ideal gas, the temperature and pressure are related by

$$C_pT = \frac{\gamma}{\gamma - 1}\frac{p}{\rho} \tag{11-117}$$

If we use Eq. (11-116) to replace ρ in Eq. (11-117), we find

$$\frac{dT}{T} = K\left(\frac{\gamma - 1}{\gamma}\right)\frac{dp}{p} \tag{11-118}$$

This can be integrated to give the relation between p and T:

$$\frac{T}{T_o} = \left(\frac{p}{p_o}\right)^{K(\gamma-1)/\gamma} \tag{11-119}$$

The pressure, temperature, flow area, and Mach number are plotted in Fig. 11-84. The Mach number is

$$M \equiv \frac{u}{c_s} = \frac{u}{\sqrt{\gamma P/\rho}} = \frac{u}{\sqrt{\gamma RT/W_m}} \tag{11-120}$$

where c_s is the local speed of sound, R is the universal gas constant, and W_m is the molecular weight of the gas. For constant velocity:

$$\frac{M}{M_o} = \left(\frac{T}{T_o}\right)^{1/2} = \left(\frac{p}{p_o}\right)^{K(\sigma-1)/2\sigma} \tag{11-121}$$

Working Fluid Conductivity. Conductivity of the working fluid is a critical parameter in a gaseous MHD topping unit. The conductivity is obtained by summing the contributions from each species. However, electrons contribute most of the conductivity due to their higher mobility.

$$\sigma = n_e e \mu_e = \frac{n_e e}{m_e \nu} \tag{11-122}$$

At low ionization states, the electron mobility is inversely related to the collision frequency with neutral particles. As the degree of ionization increases, the mobility is reduced by electron-ion and

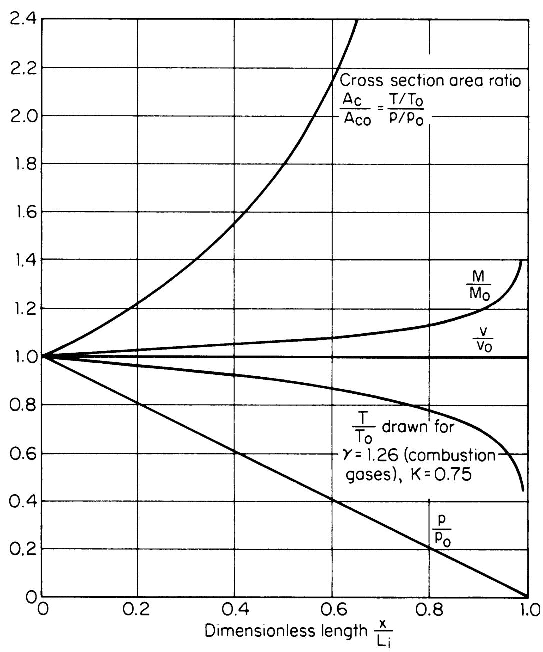


FIGURE 11-84 Dimensionless ratios vs. normalized length (x/L_i) in an ideal, constant-velocity channel.

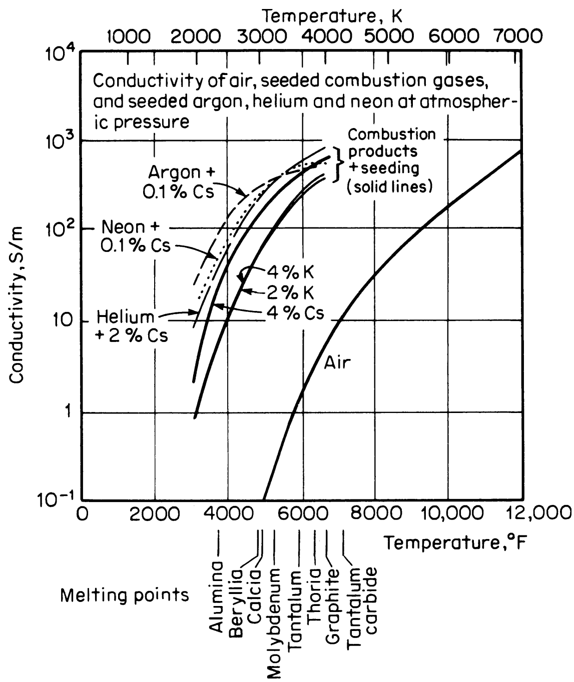


FIGURE 11-85 Conductivity of air, seeded combustion gases and seeded argon, helium, and neon at atmospheric pressure. (All curves, except for air, are from Ref. 46.)

electron-electron collisions, which have higher cross sections. At only 1% ionization, already this effect becomes significant. Therefore, a high degree of ionization is not necessary to achieve an appreciable fraction of the fully ionized conductivity.

Temperatures for fossil-fuel combustion are in the range of 2,500 to 3,000 K. At these temperatures, thermal ionization of air, combustion product gases, or inert gases is so low that the electron density is orders of magnitude below that is necessary to obtain suitable conductivity (see Fig. 11-85). One might obtain marginally acceptable conductivity by reducing the gas pressure, however, this also would result in larger duct and heat-exchanger sizes.

Fortunately, a large increase in conductivity can be obtained by seeding the gas with a small percentage of materials with much lower ionization potential. The ionization potential of the outermost electron in air is ~14 V, and that of inert gases is even higher. Alkali metals make exceptional seed materials, with ionization potential of 3.89 V for cesium, 4.34 V for potassium, and 5.4 V for lithium. Some calculated conductivities of seeded gases are shown in Fig. 11-86 as a function of temperature and pressure. In general, these curves are in close agreement with measured values (e.g. Ref.[45])

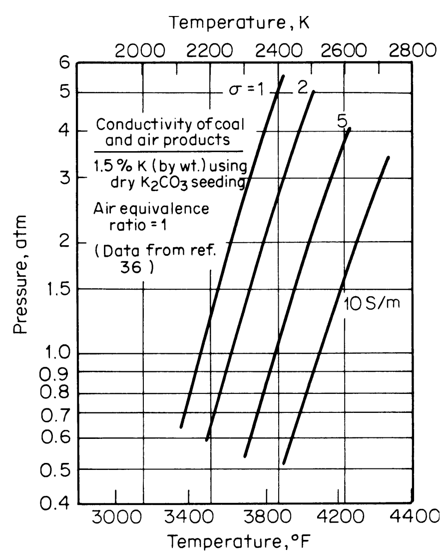


FIGURE 11-86 Conductivity vs. temperature and pressure for seeded coal products and air using 1.5 wt.% K, obtained using dry K_2CO_3. (Adapted from Ref. 46.)

Unfortunately, alkali metals also carry a relatively high collision cross section, such that an increasing percentage of seed atoms not only increases the electron density but also decreases the mobility. An optimal seeding percentage is reached at about 0.1% for cesium and potassium in argon and about 0.3% in neon.[46]

11.10.5 2-Phase MHD

Flow Characteristics. MHD of conducting 2-phase mixtures, consisting of liquid and gas phases, raises new phenomena providing the potential for unique applications. 2-Phase mixtures may arise from boiling or from mixing of distinct gas and liquid phases—for example, helium mixed with a liquid metal. Here we summarize the basic flow characteristics of 2-phase mixtures and the application to liquid metal MHD (LMMHD) for power conversion.

Much of the early progress studying 2-phase flows was based on empirical results,[47–50] since the underlying flow structures can be very complex. Similar to single-phase flows, the effect of the magnetic field is to suppress turbulence and to alter the velocity profiles. In addition, modifications in the interface configuration and slip between phases occur, and the transition between flow regimes can be shifted.

Typical 2-phase flow patterns are depicted in Fig. 11-87. As in ordinary 2-phase flow, increasing superficial gas velocity causes the flow to transition from bubbly, to churn, to slug, and finally to annular mist flow regimes.[51] However, observable differences between MHD and non-MHD behavior occur, as summarized in Fig. 11-87.

MHD Generators and Power Conversion. Liquid metal MHD power conversion using 2-phase mixtures was contemplated as early as the 1960s.[52] In the 1970s, an extensive program was conducted at Argonne National Laboratory (ANL),[50,53–56] culminating in the development of a constant-velocity dc Faraday generator using N_2 with Na or NaK. Following this early work, a rather extensive program was initiated at Ben-Gurion University, where a variety of power conversion systems have been analyzed and/or tested.[55]

The use of liquid metals for power conversion avoids the very high temperatures required to maintain an ionized gas in the conducting state. In that case, practically any heat source can be used, including solar, geothermal, nuclear, or even coal combustors. In addition, the higher conductivity of liquid metals makes possible higher power density with moderate magnetic fields, so that relatively small sized generators are possible. For example, liquid metals offer conductivities of the order of 10^6 to 10^7 $(\Omega \cdot m)^{-1}$ at low temperature, as compared with 10 $(\Omega \cdot m)^{-1}$ for the case of He seeded with 0.45% Cs at 2,000 K. Considerable support has been obtained for research on space-based power supplies, due in part to these advantages.[56]

In general, a thermodynamic cycle requires a working medium (or "thermodynamic medium") that can expand and contract with temperature, for example, gas or steam. For MHD power conversion, the thermodynamic fluid is mixed with an electrodynamic fluid (the liquid metal) to allow MHD generation.

Because the heat capacity of the liquid phase significantly exceeds the gas phase, 2-phase flow expansion (and compression) occurs nearly isothermally. This results in potentially higher thermal conversion efficiency, approaching that of the ideal Carnot cycle. For example, Fig. 11-88 compares a standard gas Brayton cycle *T-s* diagram with a modified cycle using a LMMHD generator.

Several classes of thermodynamic cycles are possible, depending on the types of coolant (e.g., gas, liquid, and/or 2-phase fluid), the use of evaporation or gas mixing, and the use of bottom cycles. These are summarized in Table 11-22.

In the "homogeneous cycle," the thermodynamic and electrodynamic fluids remain mixed throughout the cycle (Fig. 11-89). The heat source causes the working fluid to evaporate, which drives expansion through the 2-phase generator.

In an Ericsson cycle, a mixer/separator combination provides the ability to operate over wider temperature ranges. The main steps are depicted in Fig. 11-90. In this example, the top cycle has four main steps: (1) the liquid metal is heated by a heat source; (2) a mixer combines the thermodynamic fluid with the liquid metal; (3) the mixture is expanded through a generator; and (4) the two phases

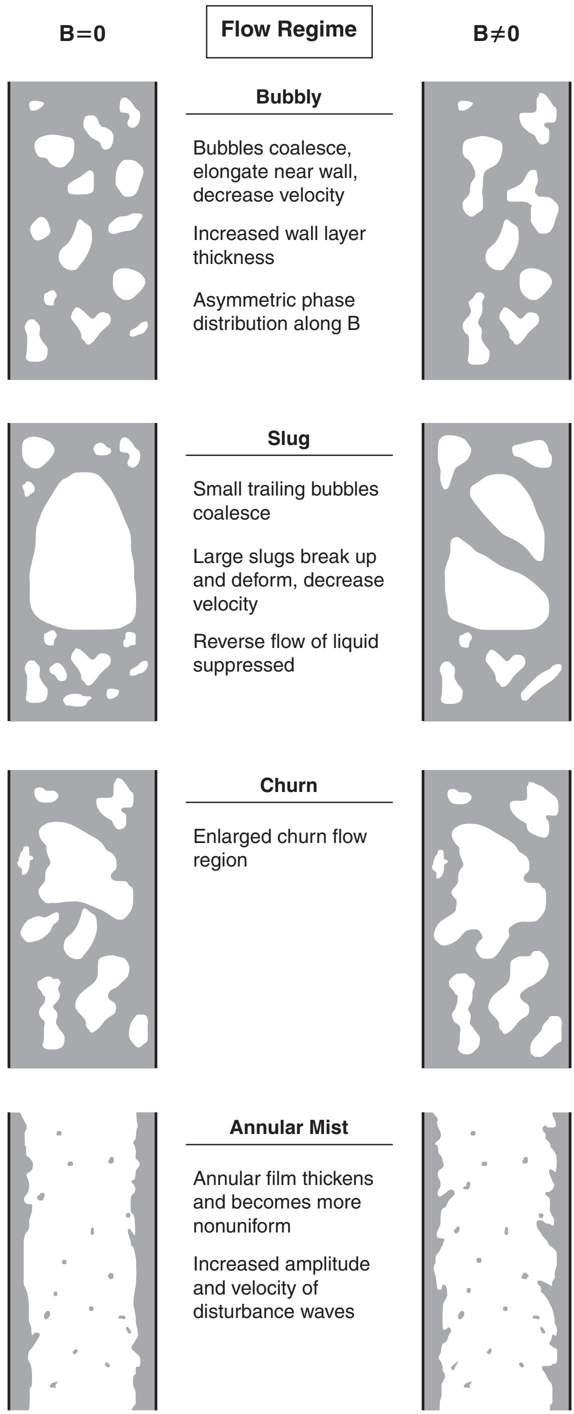


FIGURE 11-87 Effect of magnetic field on 2-phase flow regimes.

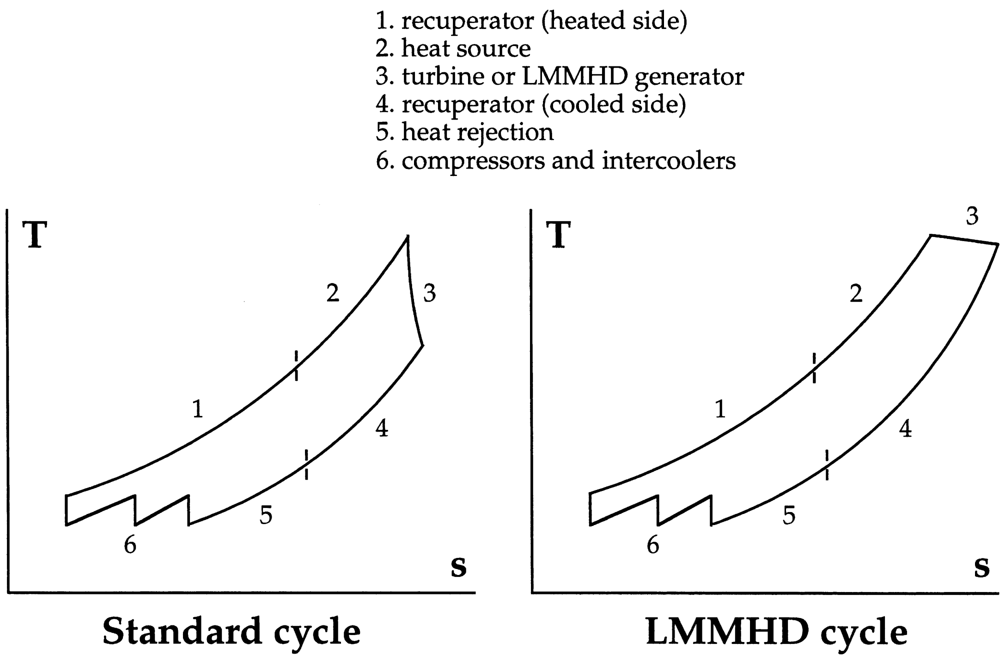


FIGURE 11-88 Standard Brayton cycle compared with MHD Brayton cycle.

are separated. The gas side of the cycle may itself take advantage of the useful heat by expansion through a Brayton cycle turbine, or it could utilize a 2-phase MHD compressor.

In the Rankine cycle, typically water is injected into a chemically compatible liquid metal (such as a lead-alloy). A steam turbine and/or 2-phase generator is used for electric generation.

11.10.6 Nomenclature

- a channel half-width parallel to B
- b channel half-width perpendicular to B
- B magnetic field intensity
- B_o characteristic field strength, used to nondimensionalize equations
- c_e RMS electron thermal velocity in a Maxwellian distribution, $(3kT_e/m_e)^{1/2}$
- c_s speed of sound
- e charge on an electron

TABLE 11-22 LMMHD Cycles and Possible Working Fluids[55]

Cycle type	Working fluids
Homogeneous cycles	Na, K, Cs, etc.
Ericsson cycle with LMMHD compressor or Brayton gas turbine/compressor	Na/He, Li/He
Rankine cycles with or without steam turbine	Pb-alloy/steam
Binary cycles (e.g., homogeneous top plus Rankine bottom)	

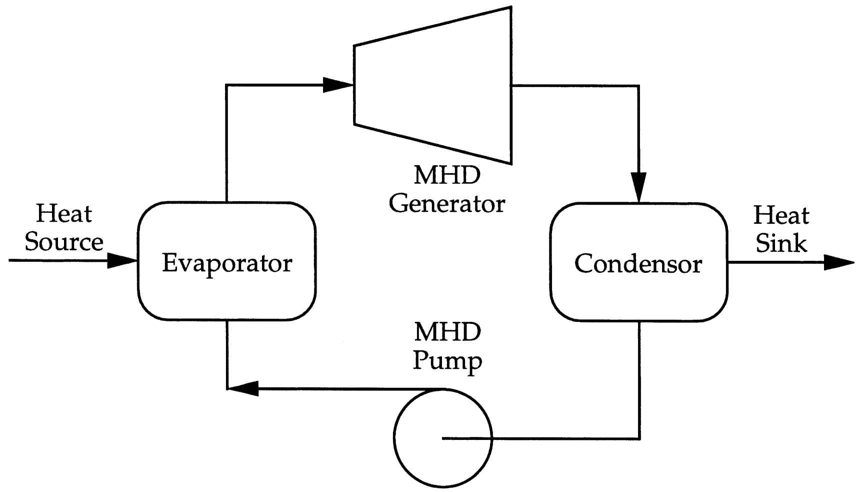


FIGURE 11-89 Homogeneous LMMHD cycle.

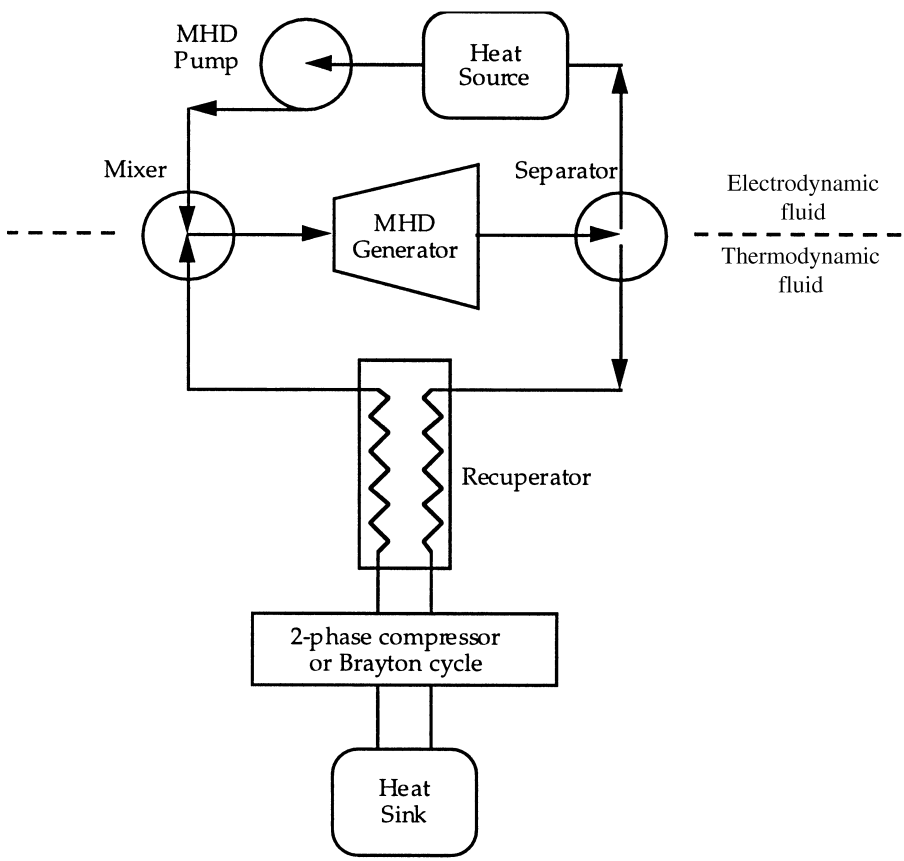


FIGURE 11-90 Modified Ericsson cycle with LM heating.

Ha	Hartmann number, $aB\sqrt{\sigma/\mu_f}$
j	current density
j_f	current density in fluid
j_w	current density in wall
k	Boltzmann constant
K	loading parameter
K_H	Hall loading parameter
l	length
m	mass
m_e	electron mass
n_e	number density of free electrons
p	fluid pressure
P	pressure gradient
q	electric charge
Re_m	magnetic Reynolds number, $\sigma\mu$vl
R_L	load resistance
T_e	electron temperature
u	fluid velocity
$\mathbf{u}_e$	drift velocity of electrons
v	particle velocity
v_b	bulk average velocity
v_o	characteristic velocity, used to non-dimensionalize equations
V	nondimensional velocity
V_{oc}	open circuit voltage
W_m	molecular weight
X	non-dimensional x coordinate
Y	non-dimensional y coordinate
β	rectangular channel aspect ratio
δ	wall thickness or skin depth
ε	electric field
λ	electron mean free path
μ_e	electron mobility, $\omega\tau/B$
μ_f	fluid dynamic viscosity
μ_m	magnetic permeability
ν	electron-atom collision frequency
ν_f	kinematic viscosity, μ_f/ρ
ρ	fluid density
σ	electrical conductivity
σ_f	electrical conductivity of fluid
σ_w	electrical conductivity of wall
τ	electron mean collision time, λ/C
φ	electric scalar potential
Φ	wall conductance ratio
ω_e	electron cyclotron frequency, $eB/m_e = \mu_e B/\tau$
ω_c	cyclotron frequency

REFERENCES

1. Northrup, E. F., "Some Newly Observed Manifestations of Forces in the Interior of an Electrical Conductor," *Phys. Rev*., vol. 24(6), p. 474, 1907.
2. Kolesnichenko, A. F., "Electromagnetic Processes in Liquid Material in the USSR and Eastern European Countries," *Iron and Steel Institute of Japan (ISIJ)*, vol. 30 (1), pp. 8–26, 1990.
3. Davidson, P. A., *An Introduction to Magnetohydrodynamics*, Cambridge University Press, 2001.

4. Sporn, P., and Kantrowitz, A., "Magnetohydrodynamics: Future Power Process?" *Power* vol. 103(11), pp. 62–65, Nov. 1959.
5. Steg, L., and Sutton, G. W., "Prospects of MHD Power Generation," *Astronautics* vol. 5, pp. 22–25, Aug. 1960.
6. Graneau, P., "Electrodynamic Seawater Jet: An Alternative to the Propeller?" *IEEE Transactions of Magnetics*, vol. 25(5), pp. 3275–3277, 1989.
7. Tsinober, A., "MHD Flow Drag Reduction," in *Viscous Drag Reduction in Boundary Layers*, American Institute of Astronoutics and Aeronautics, 1990.
8. Baker, C. C., Conn, R. W., Najmabadi, F., and Tillack, M. S., "Status and Prospects for Fusion Energy from Magnetically Confined Plasmas," *Energy*, vol. 23(7/8), pp. 649–694, 1998.
9. Spitzer, L. Jr., *Physics of Fully-Ionized Gases*, Interscience, New York, 1962.
10. Pai, S. I., *Magnetogasdynamics and Plasma Dynamics*, Springer-Verlag, Vienna, 1962.
11. Kulikovskii, A. G., and Lyubimov, G. A., *Magnetohydrodynamics*, Addison-Wesley, Reading, MA, 1965.
12. Sutton, G. W., and Sherman, A., *Engineering Magnetohydrodynamics*, McGraw-Hill, New York, 1965.
13. Moreau, R. J., *Magnetohydrodynamics*, Kluwer Academic Publishers, 1990.
14. Branover, H., *Magnetohydrodynamics Flow in Ducts*, Keter Publishing House, Jerusalem, 1978.
15. Shercliff, J. A., *A Textbook of Magnetohydrodynamics,* Pergamon Press, Oxford, 1965.
16. Shercliff, J. A., "Steady Motion of Conducting Fluids in Pipes Under Transverse Magnetic Fields," *Proc. Cambridge Philosophical Society*, vol. 49 pp. 126–144, 1953.
17. Gold, R., "Magnetohydrodynamic Pipe Flow, Part I," *J. Fluid Mechanics*, vol. 13, p. 505, 1962.
18. Chang C. C., and Lundgren, T. S., "Duct Flow in Magnetohydrodynamics," *ZAMP*, vol. 12 (2), p. 100, 1961.
19. Schercliff, J. A., "The Flow of Conducting Fluids in Circular Pipes Under Transverse Magnetic Field," J. *Fluid Mechanics*, vol. 1, p. 644, 1956.
20. Hunt, J. C. R., "Magnetohydrodynamics Flow in Rectangular Ducts," *J. Fluid Mechanics*, vol. 21(4), pp. 577–590, 1965.
21. Temperley, D. J., and Todd, L., "The Effect of Wall Conductivity in Magnetohydrodynamic Duct Flow at High Hartmann Number," *Proc. Cambridge Philosophical Society*, vol. 69, pp. 337–351, 1971.
22. Walker, J. S., "Magnetohydrodynamic Flows in Rectangular Ducts with Thin Conducting Walls," *Journal de Mecanique* vol. 20(1), pp. 79–112, 1971.
23. Alty, C. J. N, "Magnetohydrodynamic Duct Flow in a Uniform Magnetic Field of Arbitrary Orientation," *J. Fluid Mechanics*, vol. 48, p. 429, 1971.
24. Kulikovskii, A. G., "Slow Steady Flows of a Conducting Fluid at Large Hartmann Numbers," *Fluid Dynamics*, vol. 3(2), pp. 3–10, 1968.
25. Tillack, M. S., "Application of the Core Flow Approach to MHD Fluid Flow in Geometric Elements of a Fusion Reactor Blanket," in *Liquid Metal Magnetohydrodynamics*, J. Lielpeteris and R. Moreau, editors, Kluwer Academic Publishers, Dordrecht, pp. 47–53, 1989.
26. Buhler, L., "Magnetohydrodynamic Flows in Arbitrary Geometries in Strong, Nonuniform Magnetic Fields, *Fusion Technology*, vol. 27, pp. 3–24, 1995.
27. Hua, T. Q., Walker, J. S., Picologlou, B. F., and Reed, C. B., "Three-Dimensional MHD Flows in Rectangular Ducts of Liquid-Metal-Cooled Blankets," *Fusion Technology*, vol. 14, p. 1389, Nov. 1988.
28. "Permanent Magnet Flowmeter for Liquid Metal Piping Systems," Technical Report RDT C 4-5T, Division of Reactor Research and Development, U.S. Atomic Energy Commission, 1971.
29. *The Flow and Level Handbook*, vol. 28, Omega Engineering, Inc., 1992.
30. El-Wakil, M. M., *Nuclear Heat Transport*, International Textbook Company, 1982.
31. Baker, R. S., and Tessier, M. J., *Handbook of Electromagnetic Pump Technology*, Elsevier Science Publishing, New York, 1987.
32. Gilbert, J. B, and Lin, T. F., "Analyses of Underwater Magnetohydrodynamic Propulsion," Proceedings of the 25th Intersociety Energy Conversion Engineering Conference, Reno Nevada, vol. 5, IEEE 90CH2942–1, pp. 514–520, 1990.
33. Mitchell, D. L., and Gubser, D. U., "Induction-Drive Magnetohydrodynamic Propulsion," *J. Superconductivity*, vol. 6(4), pp. 227–235, 1993.
34. Murgatroyd, W. "Experiments on Magneto-Hydrodynamic Channel Flow," *Philosophical Magazine*, vol. 44,

p. 1348, 1953.

35. Henoch, C., and Stace, J., "Experimental Investigation of a Salt Water Turbulent Boundary Layer Modified by an Applied Streamwise Magnetohydrodynamic Body Force," *Physics of Fluids*, vol. 78(6), pp. 1371–1383, 1995.

36. Brooks, J. N., Allain, J. P., Bastasz, R., Doerner, R., Hassanein, A., Kaita, R., Luckhardt, S., Maingi, R., Majeski, R., Morley, N., Narula, N., Rognlien, T., Rudakov, D., Ruzic, D., Stubbers, R., Ulrickson, M., Wong, C.P.C., Ying, A., "Overview of the ALPS Program," *Fusion Science and Technology*, vol. 47(3), pp. 669–677, 2005.

37. Alpher, R. A et al., "Some Studies of Free Surface Mercury Magnetohydrodynamics," *Review of Modern Physics*, vol. 32(4), p. 758, 1960.

38. Aitov, T. N. et al., "Flow of Electrically Conducting Fluid in a Thin Layer with a Free Surface Under the Action of a Strong Magnetic Field," *Magnetohydrodynamics*, vol. 21(3), pp. 71–76, 1985.

39. Morley, N. B., and Abdou, M. A., "Study of Fully-Developed Liquid-Metal Open-Channel Flow in a Nearly Coplanar Magnetic Field," *Fusion Technology*, vol. 31, pp. 135–153, 1997.

40. Morley, N. B., Smolentsev, S., Munipalli, R., Ni, M., Gao, D., Abdou, M., "Progress on the Modeling of Liquid Metal, Free Surface, MHD Flows for Fusion Liquid Walls," *Fusion Engineering and Design*, vol. 72, pp. 3–34, 2004.

41. Hsieh, D. Y., "Stability of a Conducting Fluid Flowing Down an Inclined Plane in a Magnetic Field," P*hysics of Fluids*, vol. 8(10), 1785–1791, 1965.

42. Beinstock, D. et al., "Air Pollution Aspects of MHD Power Generation," 13th Symposium on Engineering Aspects of Magnetohydrodynamics, Stanford, CA, 1973.

43. Hals, F. A., and Lewis, P. F., "Control Techniques for Nitrogen Oxides in MHD Power Plants," 12th Symposium on Engineering Aspects of Magnetohydrodynamics, Argonne, IL, 1972.

44. Rosa, R., and Kantrowitz, A., "MHD Power," *Int. Sci. Technol*., vol. 33, Sept. 1964.

45. Brogan, T. R., "MHD Power Generation," *IEEE Spectrum*, vol. 2, Feb. 1964.

46. Frost, L. S., "Conductivity of Seeded Atmospheric Pressure Plasmas," *J. Appl. Phys*., vol. 32, Oct. 1961.

47. Owen, R. G., Hunt, J. C. R., and Collier, J. G., "Magnetohydrodynamic Pressure Drop in Ducted Two-Phase Flows," *Int. J. Multiphase Flow*, vol. 3(1), pp. 23–33, July 1976.

48. Michiyoshi, I., Funakawa, I., Kuramoto, C., Akita, Y., and Takahashi, O., "Local Properties of Vertical Mercury-Argon Two-Phase Flow in a Circular Tube under Transverse Magnetic Field," Int. *J. Multiphase Flow*, vol. 3(5), pp. 445–457, Aug. 1977.

49. Saito, M., Inoue, S., and Fujii-e, Y., "Gas-Liquid Slip Ratio and MHD Pressure Drop in Two-Phase Liquid Metal Flow in Strong Magnetic Field," *J. Nucl. Sci. Tech*., vol. 15, pp. 476–489, 1978.

50. Petrick, M. et al., "Experimental Two-Phase Liquid Metal Magnetohydrodynamic Generator Program, Final Report," ANL report MHD-79-1, 1978.

51. Serizawa, A., Ida, T., Takahshi, O., and Michiyoshi, I., "MHD Effects on NaK-Nitrogen Two-Phase Flow and Heat Transfer in a Vertical Round Tube," *Int. J. Multiphase Flow*, vol. 16(5), pp. 761–788, 1990.

52. Jackson, W. D., and Pierson, E. S., "Operating Characteristics of the M.P.D. Induction Generator," *Proc. 1st Symposium on Magnetoplasmadynamic Electric Power Generation*, Newcastle-upon- Tyne, pp. 38–42, 1962.

53. Petrick, M., and Branover, H., "Liquid Metal MHD Power Generation–its Evolution and Status," *Progress in Astronautics and Aeronautics*, American Institute of Astronautics and Aeronautics, vol. 100, pp. 371–400, 1985.

54. Fabris, G., Pierson, E. S., Pollack, I., Dauzvardis, P., and Ellis, W., "High-Power-Density Liquid-Metal MHD Generator Results," *Proc. 18th Symposium on Engineering Aspects of Magnetohydrodynamics*, Butte, MT, 1979.

55. Blumenau, L., Branover, H., El-Boher, A. Spero, E., Sukoriansky, S., Talmadge, G., and Greenspan, E., "Liquid Metal MHD Power Conversion Systems with Conventional and Nuclear Heat Sources," *Proc. 24th Symposium on Engineering Aspects of Magnetohydrodynamics*, Butte, MT, 1986.

56. Morse, F. H., "Survey of Liquid Metal Magnetohydrodynamic Energy Conversion Cycles," *Energy Conversion*, vol. 10, Pergammon Press, pp. 155–176, 1970.

SECTION 12
ELECTRIC POWER SYSTEM ECONOMICS

By Gerald B Sheblé
Honorary Distinguished Professor, Portland State University
Honorary Professor, University of Porto, Portugal
Erskine Fellow, University of Canterbury, Christchurch, New Zealand, Fellow, IEEE

CONTENTS

12.1 INTRODUCTION

The long-range trend for electric supply, demand, and costs are rising erratically and are very volatile as of this writing. The supply is changing due to recent energy deregulation legislation and also due to the rising demands for energy in a global economy of rising fuel prices. The electric demand within the United States is expected to increase dramatically.

The increase is due in part to an expected shift to hybrid or electric cars based on fuel cells, and mass transportation to replace the present fossil fuel based transportation. The need for biofuels and hydrogen fuels in diverse geographic locations will result in new and upgraded electric transmission lines for reliability and transportation of energy.

Fuel	Coal	Oil	Natural Gas	Water	Wind Bio-fuels
Transportation	Train Boat Barge Truck	Ship Pipeline	Pipeline LNG—Ship CNG—Ship, train, truck	Water Shed System—River	
Conversion	Electric Generation Plant	Electric Generation Plant	Electric Generation Plant	Electric Generation Plant	Electric Generation Plant
Transportation	Transmission Distribution		Transmission Distribution	Transmission Distribution	Distribution
Customer	Conversion to Heat, Motion, Information				

FIGURE 12-1 Leontief model of the energy industry.

A strategic infrastructure for the production and distribution of energy is essential for industrialized nations. The oil crisis of the 1970s, the first oil crisis, demonstrated the dramatic increases in the price of oil and the resulting impact on modern economies of the western nations. The recent dramatic increase in the energy demands of the Asian countries will accelerate and acerbate these crises.

A Leontief model of the energy industry is shown in Fig. 12-1. Coal was the fuel that industrialized the western countries. Oil and, now, natural gas are sustaining the western economies. Conversion of natural gas to liquid natural gas (LNG) and compressed natural gas (CNG) increases the economy of gas shipment. Hydrogen fuel will most likely start to impact the energy infrastructure in a similar way that LNG has altered the shipment of energy across the oceans. Hydrogen gas will most likely be created and used locally for the most part due to containment problems. Production of hydrogen gas by fuel cells or as a biofuel (bacteria based) is most likely to be a local or distributed process, another distributed generation plant. Hydrogen as a fuel is not yet firmly defined.

Electricity competes with direct use of fuel, such as oil and natural gas, as well as distributed generation based on wind and biofuel–based units. Energy is transported to the point of consumption either directly or primarily by pipeline, or indirectly as electricity.

An electric supply chain model is shown in Fig. 12-2. A traditional natural gas supply chain is shown in Fig. 12-3. An LNG supply chain is depicted in Fig. 12-4. Similar supply chains can be

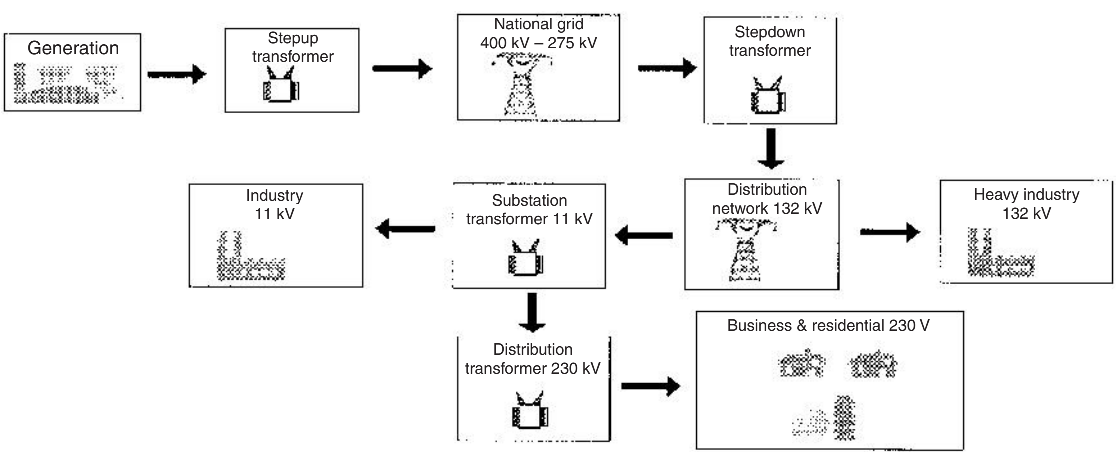


FIGURE 12-2 Electric supply chain model.

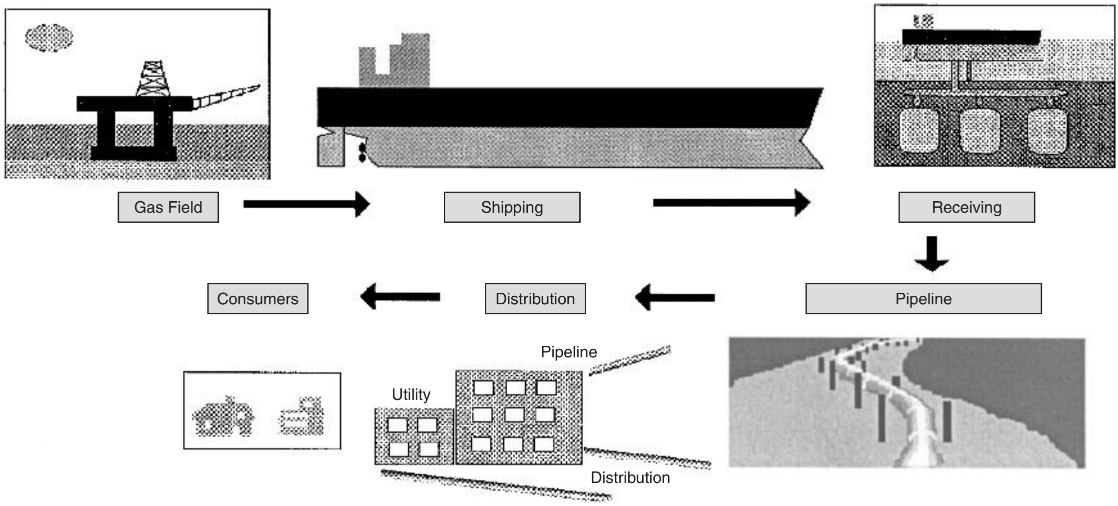


FIGURE 12-3 Natural gas supply chain.

used to depict the use of each type of fuel. Coal, for example, would include track networks for train transportation, river networks for barge transportation, and highway networks for truck transportation.

There is always the comparison of locating a generating plant near a fuel supply and transmitting the power versus locating the generation plant near the load and transporting the fuel. This is the traditional planning problem of comparing the cost of transportation by wire, train, barge, truck, or pipeline. The traditional electric system planning problem was to resolve the more economic and reliable manner of transporting fuel, such as coal, from the mine to the customer. The typical question

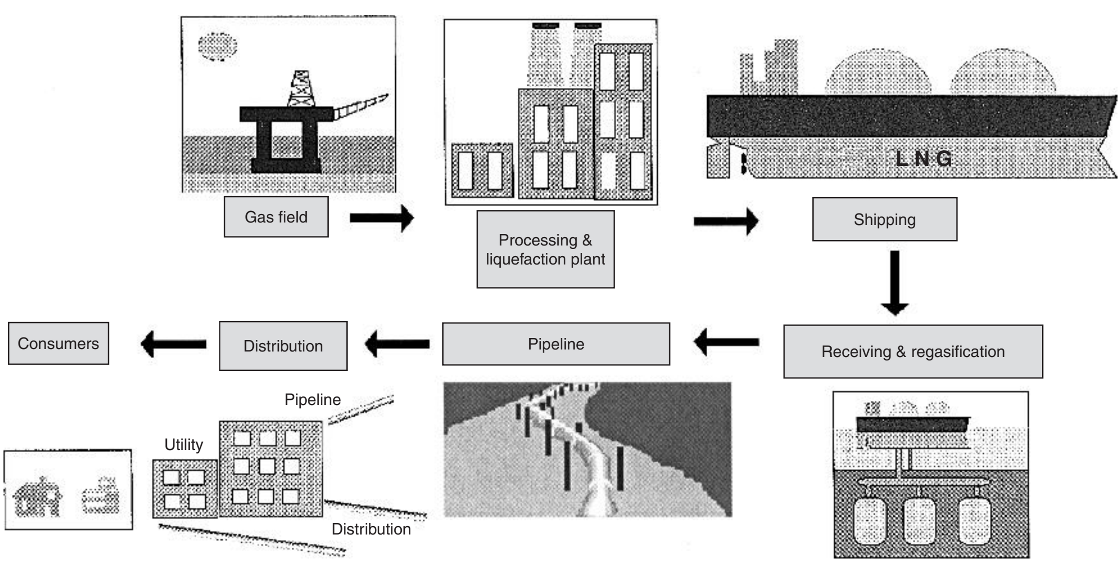


FIGURE 12-4 Liquid natural gas supply chain.

was if the *unit train* of 110 coal cars was more efficient and reliable due to the storage possibilities of coal, than an electric network, highvoltage alternating current (HVAC) or highvoltage direct current (HVDC) systems, to move the energy from the source to the customer.

The expansion of the train network for coal capacity is presently recognized as a constraint on electric system planning as this means of transportation is often congested, and suffers from restricted flows. The same question was posed in the use of oil and natural gas. The expansion of the pipeline networks has leveled the price of oil and natural gas between industrial centers. Thus, there are multiple transportation networks that are operated, maintained, and expanded to move energy from the source to the customer. Each of these networks is an energy grid.

Distributed generation has a compounding impact on the electric transmission system. Distributed generation is composed of natural gas–fired combustion turbines, biofuel–fired combustion turbines, wind generation, solar cells, and recently, gas-fired combined-cycle units consisting of combustion turbines connected directly to boilers, either solely for secondary heat conversion or for additional fuel combustion. Such distributed generation, often called renewable sources, decreases the need for the electric transmission system for basic energy delivery. Instead, the transmission system shifts to a role of providing an alternate energy source when the local distributed generation is not available to provide the desired level of reliability.

The interaction between transmission and distributed generation is complicated by the details of ownership and of interrelated operational responsibility. Many alternative or renewable energy-conversion forms are not expected to be connected to the electric grid but they do considerably alter the use of the energy grids by shifting the demand pattern. Essentially, the interrelationship of alternative resources alters the demand as the customer selects between the competing supply chains.

As a regulated industry, technological improvements reduced the cost of electricity, when all other cost factors were increasing rapidly. The cost of electric energy consists of the total delivered cost from fuel mining, fuel transportation, generation, transmission, and distribution through the supply chain. The generation cost can be broken down into three major components: fuel, equipment, and wages. The relative magnitude of these various components changes primarily in response to fuel cost changes (global economic) and environmental factors, especially due to the environmental impacts, especially as addressed by the Kyoto protocol.

By the late 1980s, the share of the total electric energy cost allocated to fuel costs had increased to 42%. This is equal to the share representing all equipment costs (i.e., generation, transmission, and distribution) at that time. This trend, however, is not expected to continue because of reduced utility dependence on oil as a primary fuel source, and as the LNG supplies increase.

Generating equipment costs have increased more than other equipment costs. In the late 1960s, annual expenditures on construction of generating equipment represented 50% of all utility construction expenditures. This share decreased to 40% by 1989. Generating equipment costs are expected to continue to increase more than other equipment costs, because of the additional costs added to power plants to accommodate environmental and other regulatory requirements. Plant costs have been rising in recent years. Table 12-1 shows average operating expenses from 1992 through 2003.

Distribution costs are determined principally by the population density of the load being served and the geographic characteristics. The shift to buried cable has significantly increased distribution costs in many countries.

Transmission was primarily needed to transport power in a regulated environment. A secondary need was to interconnect for increased reliability. In a competitive environment, more transmission is needed to remove monopoly threats and price manipulation. Transmission costs have always entered into economic comparisons of alternative generation siting. Transmission availability is a major factor when considering alternative–generation contracts in a competitive environment. Transmission limitations and costs have rendered some competitive generation sources beyond the reach of some customers.

Wages represented 26% of the total cost of electric energy in 1968. These costs decreased, compared with other costs during the 1970s. The share represented by the cost attributed to wages was about 18% in the early 1980s. The share today is decreasing as mergers and acquisitions, along with benefit reforms, such as pensions, have reduced the impact of wages for most companies.

TABLE 12-1 Average Operating Expenses for Major U.S. Investor-Owned Electric Utilities

(Miles per Kilowatthour)

Plant Type	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992
Operation												
Nuclear	8.86	8.54	8.30	8.41	8.93	9.98	11.02	9.47	9.43	9.79	10.20	10.43
Fossil Steam	2.50	2.54	2.40	2.31	2.21	2.17	2.22	2.25	2.38	2.32	2.37	2.38
Hydroelectric[1]	4.50	5.07	5.79	4.74	4.17	3 85	3 29	3.87	3.69	4 53	3 82	4 33
Gas Turbine and Small Scale[2]	2.76	2.72	3.15	4.57	5.16	3.85	4.43	5.08	3.57	4.58	6.47	10.18
Maintenance												
Nuclear	5.23	5.04	5.01	4.93	5.13	5.79	6.90	5.68	5.21	5.20	5.73	5.93
Fossil Steam	2.73	2.68	2.61	2.45	2.38	2.41	2.43	2.49	2.65	2.82	2.96	2.95
Hydroelectric[1]	3.01	3.58	3.97	2.99	2.60	2 00	2 49	2.08	2.19	2 90	2 65	3 30
Gas Turbine and Small Scale[2]	2.26	2.38	3.33	3.50	4.80	3.43	3.43	4.98	4.28	5.39	7.52	12.15
Fuel												
Nuclear	4.60	4.60	4.67	4.95	5.17	5.39	5.42	5.50	5.75	5.87	5.88	6.12
Fossil Steam	17.35	16.11	18.13	17.69	15.62	15.94	16.80	16.51	16.07	16.67	17.65	17.49
Hydroelectric[1]	—	—	--	--	--	—	--	--	—	--	—	—
Gas Turbine and Small Scale[2]	43.91	31.82	43.56	39.19	28.72	23.02	24.94	30.58	20.83	22.19	26.39	28.59
Total												
Nuclear	18.69	18.18	17.98	18.28	19.23	21.16	23.33	20.65	20.39	20.86	21.80	22.48
Fossil Steam	22.59	21.32	23.14	22.44	20.22	20.52	21.45	21.25	21.11	21.80	22.97	22.83
Hydroelectric[1]	7.51	8.65	9.76	7.73	6.77	5.86	5.78	5.95	5.89	7.43	6.47	7.63
Gas Turbine and Small Scale[2]	48.93	36.93	50.04	47.26	38.68	30.30	32.80	40.64	28.67	32.16	40.38	50.92

Two principal economic factors of bulk energy supply are the cost of the equipment and the cost of the fuel. Many combinations of these two have significant energy price impacts. Decisions between existing available and new assets can only be made after estimating all costs, including capital costs, fuel costs, wages, and maintenance costs that occur periodically.

Financial calculations are used to compare the various future scenarios of supplier and buyer interactions. Other environmental and end-use requirements, such as recreation, aesthetics, and health values, have to be included in all economic evaluations.

12.2 PRIMARY SOURCES OF ELECTRIC POWER

12.2.1 General

The primary energy sources for the production of electricity have been based on the combustion of fossil fuels (coal, oil, and natural gas) to produce steam to drive turbines. Alternatively, rivers are impounded to provide water to drive hydraulic turbines. A third principal source is the heat of nuclear reaction by uranium to produce steam to drive steam turbines.

12.2.2 Fossil Fuel Resources

During the early stages of the industrial revolution, most energy was generated by burning wood or coal in a boiler to produce steam to drive reciprocating steam engines, which, in turn, drove machinery by a system of belts and pulleys or was connected to drive wheels for locomotive use. Early electric power generation used the same process except that the belts and pulleys were connected to a generator to produce electricity. A significant advance was the development of the steam turbines.

Multiple units are generally located in one plant in order to achieve economies of scale, as common equipment can then serve more than a single unit. Common equipment includes fuel- and ash-handling equipment, water treatment, support buildings, and computer equipment, electrical equipment inventory for replacement parts, operating and maintenance staff, and transmission-line substation equipment.

Bituminous, subbituminous, and lignite are classifications given to coals to indicate the amount of heat content per measure of weight. Transportation costs are significant and thus lignite, which has the lowest heat content, is often burned only in plants located at the fuel source. Experiments to convert coal to gases have been conducted to reduce the cost of coal transportation and have been implemented with limited success.

Part of the sulfur found in coal is converted to sulfur oxides, which are considered pollutants when discharged into the atmosphere. Most of the eastern and all the midwestern coals have high sulfur content, which requires some form of sulfur-removal equipment. Such equipment significantly increases plant capital costs and reduces plant efficiency. Coals with lower sulfur content are located in some western states. Transportation costs to bring this coal to the East and Middle West add significantly to its cost.

Presently, it is not financially feasible to convert coal to gaseous or liquid fuel, but it is an area of increased research and development. These procedures are attractive because they offer the possibilities of sulfur removal before combustion and of providing fuel for combustion turbines as well as steam boilers.

Boilers and precipitators are designed for the specific heat content, and so and so other physical and chemical properties (like sulfur content) of the fuel need to be used. Rising fuel costs have justified the conversion of many units to the use of multiple fuels.

Biofuels are quickly coming to the forefront as the price of oil escalates. Bio-fuels include ethanol, soy diesel, and gases produced from agricultural sources and animal and human wastes. Recent advances in waste processing have led to the building of power plants in conjunction with waste treatment, especially the waste from animal herds. Plans have been announced to convert human wastes into gases in the near future, partly as a response to reduce the environmental impact of waste-water treatment.

Solid waste is currently being used as a fuel and as an additive to coal in conventional power plants. Such combination fuel burning was in response to landfill limitations, but the increasing cost of oil is starting to justify the active use of waste resources.

Combustion turbines use gaseous and liquid fossil fuels that are burned, such that the hot gases can be used to drive a turbine directly. These combustion turbines eliminate the conversion of energy to steam and subsequent conversion to electricity, and thus have lower costs due to this system reduction. Such combustion turbines are less efficient and require more expensive fuels and more maintenance. The net economic impact is higher operating costs. Recent developments have increased their efficiencies significantly by using the exhaust output of several units as input to a boiler system to create steam as a traditional unit performs. The output of a combustion turbine is a high heat content exhaust gas. Not only is this gas at a high temperature, but it also contains a considerable amount of unburned fuel. It is economically possible to use the exhaust gas to generate steam either directly in a waste-heat recovery boiler or as preheated combustion air into a conventional boiler with the addition of other fuels. The steam produced can then drive a steam turbine-generator.

This arrangement is called a *combined-cycle plant*. Internal combustion engines are used to drive electric generators at distributed sites for reliability of supply. Hospitals, airports, emergency facilities, communication facilities, and other infrastructure needs require distributed generation to achieve significantly increased reliability requirements. Due to the operating costs of such facilities, they do not represent a significant part of total power generation at this time.

Residual fuel oil is a significant source of energy for power production. This oil contains the heavier components of crude oil that remain after gasoline and other light hydrocarbons have been removed. Oil-fired steam power plants are less expensive to build and operate than coal-fired plants. Combustion turbines use lighter oils as fuel.

Natural gas was traditionally a fuel for steam power plants located near oil fields where the gas is produced. The clean burning properties of gas have lead to gas firing in coal or oil boilers in other parts of the country as natural gas pipeline capacity is available. As there is a high value of natural gas for chemical and space-heating uses, its future use as an energy source for electric generation is limited. Natural gas is a significant fuel for distributed generation, especially if the heat can be used locally. Such generation includes combustion turbines that readily use natural gas as a fuel, especially when combined with an additional heat recovery system, called a combined–cycle plant.

12.2.3 Nuclear Fuel

Nuclear reactors were developed as economical electric power production when the long–term storage of the spent nuclear fuel was considered inexpensive. Subsequent studies have led to a mixed conclusion as to whether or not spent fuel can be stored economically over the lifetime of the radioactivity.

No new nuclear units have been built recently in the United States due to the concerns of long–term storage and potential run-away reactions. Many other countries have continued to develop nuclear energy, given the increasing shortage of fossil fuels. The long–term storage of spent fuel continues to be an unsolved problem. Several existing facilities have presently reached the maximum local storage of spent fuel. Additionally, many existing units in the United States are approaching the end of their useful life cycles.

Natural uranium is the basic fuel for all heavy-water fission reactors. It must be enriched (the content of fissionable uranium increased from the natural value of about 0.7% to about 3%) to be usable. This increase is accomplished by passing the natural product through filters or centrifuges that increase the concentration of fissionable material in part of the output while reducing it in the remainder, which is then unusable as fuel. A breeder reactor converts this depleted uranium back into usable fuel, thereby greatly extending the amount of usable uranium.

Plutonium is a fissionable by-product of nuclear reactor operation. It can be mixed into natural or enriched uranium to recover the energy available in the plutonium.

Fusion reactors are expected to use deuterium as fuel. This material exists in large quantities in water but would have to be extracted and converted into a usable form as is presently under investigation as a multinational experiment.

12.2.4 Hydroelectric Power

Natural precipitation as rain or snow provides a continuous source of water at elevations higher than sea level. The flow of water back to lower elevations provides a source of energy by converting the potential energy into kinetic energy using waterwheels. Impoundment of rivers by dams provides a steady energy source and a larger elevation difference to localize the potential energy. The higher elevation of water locally is measured as effective water head. The natural elevation differential Niagara Falls was used for the motive power for the first commercial alternating current (ac) central station.

Hydropower is a renewable fuel resource. However, the traditional harnessing of hydropower is complicated by the need to dedicate a significant part of a river course to form a lake large enough to provide a steady water source. Initial costs for the dam and other construction work are significantly higher than for other types of generation. This higher first cost must be offset by long-time fuel cost savings. Therefore, the justification of hydropower is very sensitive to the replacement of other fuels and the scheduling procedures. Often the cost of a project is divided between multiple uses of water, such as power, navigation, irrigation, recreation, and flood control. These competing uses greatly restrict the availability (and thus the relative cost) of the power.

The use of tidal movement of water to generate power has been proposed in some coastal locations where there are large tides. Because of the relatively low water head provided by tidal action, it was originally thought necessary to impound huge quantities of water. The cost of the impounding structures has been found to be prohibitive. The structures also probably would have a significant environmental impact owing to their great size. A new alternative is to use wind generators to harness the energy in tidal, river, and ocean currents. Such water generators resemble wind generators but are inverted, suspended from the surface, restricted to locational movement by anchors, and spin as the current flows across the blades, roughly at the speed of a revolving door. While navigational use of that immediate area is restricted, the impact is considerably different from conventional hydro facilities. It is expected that the low cost of such systems may revitalize many of the abandoned hydro facilities with low head capability.

Efforts are being made to develop power from ocean-wave action, but these are experimental and have a significant impact on the aesthetic shoreline use.

12.2.5 Geothermal Steam

At several locations in the world, natural steam is close enough to the surface of the earth that is accessible by using conventional drilling methods to pipe it to the surface. These locations are too few to be of any overall significance to most countries. The expansion of the use of geothermal steam to areas where the heat is not near the surface will require major progress in the development of very deep–well drilling technology. There is a considerable cost to the maintenance of such units as the steam has significant quantities of corrosive and solid materials that reduce the life-expectancy of heat transfer equipment. However, the availability of such steam in several locations could be harnessed to generate hydrogen within the near future for export to energy–dependent regions.

12.2.6 Fuel Cells

Fuel cells generate low-level direct-current (dc) power as a result of a chemical reaction between a hydrocarbon fuel and oxygen. Development has progressed to the point where practical devices are available, even with the use of natural gas and other biogases. However, the costs have not yet been reduced to the point where fuel cells can be considered as competitive with other conventional power sources except in special applications where highly reliable power sources are required or in remote locations.

12.2.7 Primary Batteries

Primary batteries use a chemical reaction between two components of the battery to produce dc power. The battery components are depleted up in the process. At present, the cost is prohibitive for large-scale applications.

12.2.8 Solar Electric Power

Electric power can be developed from the sun's rays in two ways: solar cells that produce low levels of dc power as a result of the sun's rays striking certain materials and solar boilers that consist of a system of mirrors that concentrate the rays from a large area onto a vessel containing water. Practical use of solar electric power must overcome two fundamental problems: (1) the sun's energy is so diffuse that very large earth surface areas must be covered by the mechanism used to collect and convert the energy; and (2) practical energy output is limited to part of the daylight hours on cloudless days. The practical locations in the United States are in the southwestern deserts, which are relatively far from power-consuming areas as to require major transmission lines to deliver the power.

The diffuse nature of the sunlight can be harvested by the use of many photovoltaic solar cells located on all homes within a regional area. Several home owners in the southwest part of the United States have invested in such systems as the price of oil has risen significantly. Presently, such installation cost in the range of $15,000 to $20,000. The net profit from such installations was demonstrated as $200 to $300 per month in 2005. Home owners were pleased with this return on investment while reducing the dependency of the country on oil.

The use of solar collectors to power a conventional boiler have been constructed and demonstrated. Maintenance costs are high as the reflective surfaces are easily contaminated and abraded in such environments. Research into more resistant materials may soon render it possible to justify the conversion of solar energy to steam, solely on the basis of the fuel that is not consumed.

12.2.9 Wind Power

It is practical to generate power from propeller-driven generators. Recent developments in the capability of equipment and the advanced controls to cope with the variable nature of the wind and demand have lead to a major shift to use wind as a primary source of electricity. The European Union and several U.S. investors have committed to major wind development investments. Several European countries have shifted to a high penetration of wind generation, as high as 61% in the Netherlands, due to expected scarcity of fossil fuels in their regions. Costs have been significantly reduced, while the equipment reliability has been dramatically improved. The use of wind generation is easily justified for remote areas.

12.2.10 Distributed Generation

There are two generic types of distributed generation. Distributed generation is inherent when renewable resources are the fuel, such as biofuels, solar, and wind. Distributed generation is also justified when heat or steam can serve other uses.

Several companies have developed small gas-fired generating units that are intended to be located in small groups scattered throughout the distribution system. The first units were designed to be 50 kW in size. Such systems have been installed and justified when the heat is also used for environmental heating or manufacturing processes. In remote locations where electrical systems do not exist, it is expected that one or two extra units can be installed if biofuels are available. Such systems are used extensively as backup and for unplanned expansion. Many of these are operated as stand-alone systems.

It is necessary to have alternative power sources to supply the load when the sun is not shining, the water flow is reduced in dry season, and the wind is not blowing at the proper speed. These alternative resources include energy storage and demand–side management, as well as the use of conventional power plants. Thus, many of the renewable energy systems (wind, water, biofuels, etc.) require alternative sources, such as conventional power systems, or local storage. Local storage can include heat storage as well as hydrogen–based fuel cells.

There are industrial processes that require large amounts of heat at temperatures and pressures below those at which boilers generate steam. When such combined demands are served by an integrated power plant, it is possible to obtain low–cost power by generating steam at a higher temperature and pressure and running it through a turbine, subsequently exhausting the steam in the condition required by the industrial process. This arrangement for multiple uses is called a cogeneration unit (traditionally called a topping unit) as such a joint service capability provides economical energy for the following reasons:

1. The additional construction cost for the higher-temperature and higher-pressure boiler plant is not significantly higher than the cost of a boiler plant built to supply the industrial process demand only.
2. The required additional fuel generating higher-temperature and higher-pressure steam is less than the fuel cost for generating steam for industrial demand only. One principal reason for this is that for a conventional generating unit, the steam must be condensed back into water to obtain good overall efficiency. The condenser used for this purpose must be supplied with cooling water that absorbs most of the heat in the steam exhausted from the turbine. This heat is then dissipates to the atmosphere. There is far less condenser heat loss because the exhaust steam is used for process heat.
3. Cogeneration units are installed in many facilities requiring high reliability for industrial processes.

Another method of producing by-product power is the use of an extraction turbine, which has openings at one or more points to allow steam to be removed after it has passed partway through the turbine. This steam is at a lower temperature and pressure than the inlet steam and can be used as process steam. As with a topping unit, the extraction steam does not lose heat to a condenser; therefore, its generation efficiency is very high.

12.3 ENERGY STORAGE SYSTEMS

12.3.1 General Aspects

Electric power is a highly perishable commodity. There is no means of storing it directly in an electrical form. Thus, sufficient generating capacity must be constructed to meet the peak load. This expensive capacity is underused during off-peak periods. Energy-storage systems can reduce the overall cost of power by reducing the amount of generating capacity required. The storage system absorbs energy during off-peak periods and delivers it to the load during peak periods. To be economically effective, the storage system's construction cost must be low and its efficiency high.

12.3.2 Pumped-Storage Hydro

Storing energy can be accomplished by using an electric motor-driven pump to raise water from a lower pool to an upper reservoir when the electric load demand is low (at night or on weekends) or when excess generating capacity is available. Later, the same motor pump can be operated in reverse as a turbine-generator using the water in the upper reservoir as an energy source.

For a pumped-storage system to be economically justified, the power-source fuel cost must be very low (hydro, nuclear, high-efficiency fossil, solar), and the construction cost of the pumped-storage plant must be lower than alternative generating capacity. Low construction costs per unit of power require a very large capacity plant and a large elevation difference between the upper and lower pools (doubling the water head cuts the required storage volume in half). There are not many locations where the topography is suitable for this type of installation.

12.3.3 Hydrogen Fuel Cycle

A scheme that has been proposed to store the energy output of low-fuel-cost plants when they are not required to supply load is the hydrogen fuel cycle. The surplus generating capacity would be used to obtain hydrogen from water by electrolysis. The hydrogen would then be stored or transported for use as a fuel in another generating unit. Much development work will be required to determine the overall costs for this system.

12.3.4 Storage Batteries

Practical storage-battery systems are available to store surplus electrical energy in chemical form for use at a later time. However, at the present time, overall cost benefits have not been sufficient to justify the large-scale trial installations that are needed to verify costs and reliability. Research has instead been conducted on fuel cells that serve the equivalent purpose.

12.3.5 Cryogenic Storage Magnets

Research has been conducted on large cryogenic (supercold) magnets that have the capability of storing large amounts of energy in their magnetic field for long periods of time because of the very low electrical losses in the magnet conductors. Much additional research and development are required before the relative economics of this device can be determined.

12.3.6 Flywheels

The use of mechanical flywheels has been proposed for energy storage. Major development of strong materials will have to be made and pilot plants built to demonstrate the reliability and costs for this type of storage before it can be justified. Funding for this research has been reduced.

12.4 DEVELOPMENT OF ELECTRIC POWER SYSTEMS

12.4.1 Need for Fuel, Demand, and Price Forecast

The process of deregulating the electric industry is still very much ongoing. There are still many questions that haven't been answered regarding how the markets should operate and what is an appropriate market design. The instability of some electric markets has affected other industries, such as the fuel industry. As previously stated, the national load decreased after the 1973 oil crisis. With these industries so closely tied together, it has become harder to provide an accurate forecast on load growth. With the fuel markets seeing record high prices, will demand respond to the price hike and drop? To what extend would it drop? Would the drop be temporary? Could this actually

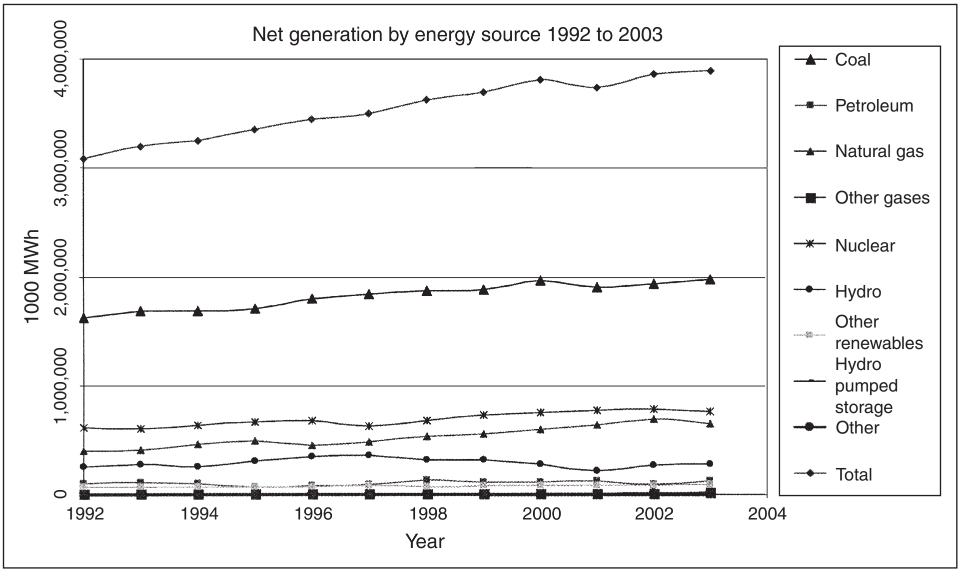


FIGURE 12-5 Net generation by energy source in the U.S.

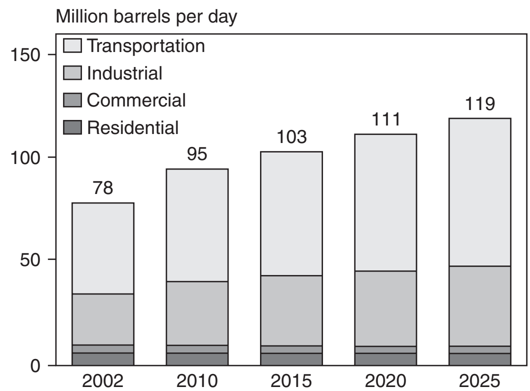
cause a spike in load growth due to products being switched from fuels to electric loads as we are seeing with hybrid cars? These are all important to understand.

Figure 12-5 shows the net generation in United States from 1992 to 2003 by energy source. The top line is the total generation. Notice that the only decrease in total generation was right after the California market crisis, though the decrease was temporary.

The total net generation has grown just over 2% a year. The fuel types that have the largest growth is natural gas and other (i.e., wind, solar, etc.) with growth rates over 4%. The high growth rate in other, which is mainly renewables like wind and solar, describes the market's concern over the dependency on oil as well as environmental implications. These sources are also becoming more economically justifiable when compared to fossil fuels due to the increase in fossil fuel costs. This concern to invest in energy sources that are more expensive is a sign of how the buyers are beginning to affect the deregulated markets now. Under regulation, there would be no high–price spikes, so assuming the demand to be inelastic was considered not to be an issue.

With the deregulated markets and the frequent market prices, this assumption is becoming more invalid. Though buyers still do not have a direct choice on who they can buy their energy from yet, it is still apparent that the views of the buyers are affecting the markets based on the large growth in renewables. As of right now, the buying side of the electric industry is not deregulated but many support such an idea. If and when that happens, people's preferences as to how their electricity is generated (i.e., environmental impacts) will come into play and those doing the forecasts right now have never dealt with this before; thus, predicting what will happen is a challenge. There is also the suggested market setup where buyers who want better reliability can pay for such and those less dependent on having a reliable connection can pay less. The markets are essentially including more of the buyers' preferences and this will only complicate the forecasting.

Right now most car dealerships have a waiting list for hybrids since they are so popular and this has caused a high demand spike in the industries that create the battery packs. In return, the efficiency in the battery packs is improving due to investment possibilities as well as the costs are decreasing. Figure 12-6 shows the distribution of oil consumption by the different sectors. With transportation being the largest sector and having the highest growth rate, it is easy to understand how even a small conversion from oil to electrical loads can cause predictions to be off.



Sources: History: Energy information administration (EIA), *International Energy Annual 2002*, DOE/EIA-0219(2002) (Washington, DC, March 2004), website www.eia.doe.gov/iea/. Projections: EIA, system for the analysis of global energy markets (2005).

FIGURE 12-6 World oil production by end-use sector.

With the growth of load very dependent on fuel costs, another issue that is arising comes from the emerging economies like China that are greatly increasing their fuel and energy consumption and thus causing a large growth in demand for the entire world. This factor is probably the most influential with regard to load–growth forecasts since predicting what will happen with these foreign countries complicates it immensely instead of just focusing on what is happening within the United States.

It is predicted that the oil consumption in emerging Asia will double between 2002 and 2025. With such strong changes in an industry as this, forecasting becomes very difficult.

Likewise, updating a forecast to be able to account for the exercise of market power is extremely difficult. The Organization of the Petroleum Exporting Countries (OPEC) share of the world oil production market is predicted to increase. Thus, their market power will increase as well. The mature economies are predicted to have very low growth in coal consumption. This reflects the demand side pressure for fuels that are less harmful to the environment.

The cost of fuel is dependent on a lot of variables: supply, demand, market power, etc. There is expected to be a large jump in the petroleum and natural gas markets while the cost of coal is relatively unchanged. The main reason is, though the cost of coal is cheaper, the United States is not investing a lot in coal due to political, environmental impacts, etc.

The demand for oil will increase; however, the price of oil will be more dependent on political, economical, and environmental concerns. These concerns will be the driving forces in determining whether prices will be high or not, since it is predicted that there will not be a problematic scarcity of oil through 2025. These concerns range from the influence governments have on those controlling the markets like OPEC to future environmental laws that might be enacted.

12.4.2 Basic Market Economic Concepts

The major change of industry restructuring is the use of markets to connect the information needs between each link in the supply chain instead of a separate company in a power geographic area. Several markets are implemented to properly price the necessary services to produce and to transport electric energy at an acceptable level of availability. The spot market is for immediate delivery of the commodity. The forward market is for the near–term delivery of the commodity. The future markets are for the financial hedging of the commodity. The bilateral markets include all contracts not traded but executed and committed for commodity delivery. The contingent markets are for potential trades of a commodity under unusual operating conditions, such as spinning reserves. The option markets include buyer selection contracts that give the right but not the requirement to execute a commodity–exchange contract. Energy balancing and load following markets are used to match the periodic (hourly) energy contracts with the actual instantaneous demand.

A market's basic purpose is to enable buyers and sellers to determine the best allocation of resources which investment is optimal, and so on. Additional commodities, originally organized by the U.S. government, are emission rights that can be traded. Such trading has been a core component of the Kyoto Agreement that is yet to be implemented but has been signed by most of the world governments.

12.4.3 Capital Budgeting Financial Economics

The basic criterion generally used in regulated electric power economic analyses is that the alternative requiring least revenue from the customer is the proper economic choice (Jeynes, 1968). The basis for re-regulated economic analysis is financial analysis as applied to any financial instrument or commodity. The essential process is to determine the costs of investment, costs of operation and of maintenance, and then to compare these costs to the expected revenue over a time horizon. The demand and supply forecasts are discussed in the previous section.

The cost of capital is the major fixed charge associated with an investment, as it is necessary to determine the total payments needed to pay the bonds, financial instruments (mortgage), and stock dividends associated with the investment cost of that equipment. Labor and material cost, including overhead and profit of the supplier (direct construction cost), is easier to track as the facilities are installed. The use of activity–based costs (ABC Accounting) is used extensively to track expenses per task for proper assignment of cost factors.

The cost of fuel includes the price of removing the raw material at the mine or well, processing costs, and the cost of the transportation system (railroad, pipeline, shipping, etc.). Usually, the cost of fuel is reduced to a single figure expressed in dollars per million British thermal units (Btus) as demonstrated by the NYMEX futures contracts. The expense to process and deliver the material is the next cost component to be included. The cost of fuel is priced, based on other fuel market prices due to value of energy selected by the consumer. Increases in the cost of production due to environmental regulations, increases in labor costs, and increases in transportation costs are the components with the largest increases as of this writing.

The amount of fuel in inventory is usually the amount required for 2 or 3 month's operation for each plant. For nuclear fuel, the carrying charges are very high as costs for expended fuel maintenance is increasing as storage costs are clouded by political uncertainty. Generating unit efficiency is stated in terms of the heat content (Btus) required from fuel to produce electric energy (kilowatthour). This conversion curve shows the quantity of fuel energy converted to electric energy (Btu/kWh) is called the *heat rate.* The net cost of fuel (dollars/kWh) is the product of the raw-fuel cost (dollars/million Btus) and unit heat rate (Btu/kWh).

Operation and maintenance costs consist of the labor expenses, including overheads for the plant operating and maintenance personnel, and maintenance and operating materials (other than fuel). These costs are modeled as a fixed component (dollars per year) and a variable component that varies with the amount of power produced (dollars/MWh). Since the total of the two is a smaller value compared with ownership and fuel costs, it is common to average the two components into a single figure. A figure of about $15/kW per year is the median value, with many plants within $\pm 50\%$ of this value and a few beyond that range. Within this wide range there is no correlation with unit size, type (coal, oil, nuclear), or number of units in the plant as the financial economics demonstrate that costs beyond that range are not viable projects.

12.4.4 Financial Engineering Methods of Analysis

Comparison of alternative means of providing power requires combining costs that occur at different times. The cost for constructing a plant occurs over a several-year period prior to initial operation of the facility. Financial analysis is based on the general concepts of capital asset pricing theory, alternatively, arbitrage–pricing theory. Once the future costs of fuel, construction, operation, and maintenance are forecast, then the demand is forecasted to determine the probabilistic revenue that could be obtained by those assets. The basic consideration is that the expected profit has to be sufficient to pay the risk premium for the expected relative corporate risk.

BIBLIOGRAPHY

Baxter, M. 1996. *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge University Press.

Best, R. 2004. *Market-Based Management***.** 4th ed. Prentice Hall.

Bodie, Z., and R. C. Merton. 1999. *Finance.* Prentice Hall.

Electric Power Annual. Washington: Energy Information Administration. (http://www.doe.eia.gov.)

Hull, J. C. 2002. *Options, Futures, and Other Derivatives.* 5th ed. Prentice Hall.

International Energy Outlook. 2005. Washington: Energy Information Administration. (http://www.doe.eia.gov).

Luenberger, D. 1998. *Investment Science*. Oxford University Press.

Neftci, S. N. 2002. *Introduction to the Mathematics of Financial Derivatives*. 2nd ed. Academic Press.

Neftci, S. N. *Principles of Financial Engineering*. Academic Press.

Statistical Yearbook of the Electric Utility Industry. Washington: Edison Electric Institute.

SECTION 13
PROJECT ECONOMICS

Allen L. Clapp
President, Clapp Research Associates, P.C., Member, IEEE

CONTENTS

13.1 BOTTOM-LINE ECONOMIC MEASUREMENTS

This primer is intended to give a quick introduction to the financial considerations that drive the decisions to start or abandon a project. The *bottom line* on any project is that it is either better or worse than alternative investments. Money is the usual medium for measuring "better" because all the other factors like risk, reputation, and enjoyment, often can be translated into a monetary equivalent.

The decision to start a project, and the selection of the method to finance it, may involve many interrelated factors. Chief among these factors are the values of project costs and receipts, interest rates, possible returns from other projects, tax regulations, and available financing. The remainder of this primer briefly discusses these items and illustrates the economic differences resulting from three different methods of financing a project: (1) 100% financing by the owner, (2) 50% owner's equity and 50% borrowed debt, and (3) leasing from another owner.

The illustrations herein are intended to convey the certain knowledge that taking shortcuts on economic analysis may lead to an inappropriate decision. This is particularly true when a long-term project, like a new energy production system, is being evaluated against a short-term project, like purchasing specialty machinery for producing a product which has a limited sales life. The correct decision is the one which yields the greatest total value to the owner.

13.2 THE VALUE OF MONEY

Money has no value of its own; its value is proportional only to the goods and services it provides. The amount of goods and services money can provide in a given year relates directly to the relative value of money at that one point in time. If inflation did not reduce the value of money over time, a specified amount of dollars could buy the same set of goods and services in one time as in another. Because of inflation, however, the value of that specific amount of dollars decreases over time; the same amount of money is worth fewer goods and services in later periods. As a result, the decision to start a project should consider both the *amounts* of expenditures and receipts associated with the project and the *timing* of those cash flows.

TABLE 13-1 Relationship of Nominal Dollars to Real Dollars

	Dollars in year of receipt					
	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Nominal value (actual dollars)	100.00	100.00	100.00	100.00	100.00	500.00
Real value (1984 buying power)	100.00	90.91	82.64	75.13	68.30	416.98
Nominal value (actual dollars)	100.00	110.00	121.00	133.10	146.41	610.51
Real value (1984 buying power)	100.00	100.00	100.00	100.00	100.00	500.00

The terms used to express the effect of time on the value of money are (1) *real dollars* and (2) *nominal-year dollars* or *nominal dollars*. Nominal dollars refer to the amount of dollars received or spent in a given year. Because of inflation, a dollar received in year *X* will be worth more or less than a dollar received in year *Y*. In order to compare the two, the real purchasing power of a year *X* dollar must be compared against the real purchasing power of a year *Y* dollar. It makes no difference whether (1) year *X* dollars are converted into the number of year *Y* dollars that have the same real purchasing power, or (2) year *Y* dollars are converted into year *X* dollars. If more convenient, both can be converted into equivalent dollars of some other nominal year.

In order to consider the effects of inflation on a project, all cash flows from each of the various years of the project should be expressed on a directly comparable, common year basis so that their relative values can be considered. To accomplish this, the *nominal-year dollars* of cash flow in each future year are converted to *constant-year dollars* by *discounting* their value back to that of one common year.

If inflation is running at 10% per year, the relative value of $100 in hand in year 1 will be $110 in year 2 or $121 in year 3, etc. Likewise, future values must be discounted to obtain their value today. In other words, the *real value* of $146.41 (nominal-year dollars) received 4 years away is only $100.00 in year 1 dollars. The illustration in Table 13-1 uses such a 10% *discount rate* to calculate the real value (in constant year 1 dollars) of future nominal-year dollars for each year. The first two rows show the decline in real value (the ability to purchase goods and services) of a stream of $100 annual receipts. The second two rows show the increase in annual dollar receipts required to maintain the same real income in each future year.

If a project is to be a success, the sum of its real costs and real returns must be positive enough to overcome any uncertainty about the occurrence of future costs and returns. The *present value* of a future income stream (or cost stream) is the sum of the *real* values of the individual future receipts (or costs). The *net present value* (NPV) of a project is calculated by subtracting the present value of project costs from the present value of expected project returns. The example in Table 13-2 illustrates both the time value of money and the process of calculating the net present value of a project. The nominal dollar values of cash stream A are identical, but in reverse order, to those in cash stream B.

In this illustration, if B is a revenue stream and A is a cost stream, the project makes some money in 5 years; the NPV is a positive value of $30. If only the nominal dollar flows are considered, the project appears to break even; the *nominal return* over the life of the project is zero. However, that

TABLE 13-2 Calculation of Net Present Value

	Yearly cash flow, $								Total cash flow, $	
	Year 1		Year 2		Year 3		Year 4			
Cash stream	Nom.*	NPV†	Nom.	NPV	Nom.	NPV	Nom.	NPV	Nominal value	Present value @ 10% discount rate
A	100	91	150	124	180	135	220	150	650	500
B	220	200	180	149	150	113	100	68	650	530
B – A	120	109	30	25	−30	−22	−120	−82	0	30

*Nom. = nominal value.
†NPV = net present value.

is *not* the case in real terms. Because of the time difference in the cash flows, the project earns a net positive real spendable return.

In this example, the project begins to lose money in year 3. Obviously, if the project can be stopped at the appropriate time, more income will be retained by the owner. If not, the project may still be the best alternative, especially if the scenario of Table 13-2 is the worst expected case and the "best guess" case would return significant profits. Whether this particular project would be started depends on such factors as the relative returns that can be earned from alternative projects, the relative risk of each project, the availability of financing, and the type and usefulness of tax advantages.

Annual Charges. It is desirable to have a convenient method of calculating the annual costs of capital investments made in an alternative scheme. Fortunately, this often can be done realistically by using a level carrying charge which is expressed as a percentage of the original investment.

The total revenue requirements of a piece of equipment are the sum of the annual charges for

1. Return on investment
2. Depreciation
3. Income tax
4. Property taxes
5. Insurance
6. Operating and maintenance expenses

The first five of these charges can conveniently be estimated as a percentage of original investment. The operating and maintenance charges should be estimated separately for each project because they do not relate to capital investment as a percentage.

Level Annual Carrying Charges. The level annual carrying charge is the percentage by which the capital investment can be multiplied to determine its annual cost on a uniform basis. The value of this carrying charge is very much dependent on the expected life of the piece of equipment because depreciation varies in accordance with life expectancy. A method of obtaining the level annual carrying charge is as follows: (1) calculate the sum of the annual charges for return on investment, depreciation, income tax, property tax, and insurance for each year of the expected life of the piece of equipment, (2) use the appropriate present-worth factor with each annual cost to convert the annual cost to a present-worth value; (3) sum up these values to obtain the total present worth of the annual carrying charges; and (4) multiply the total present worth by the capital recovery factor (see Fig. 13-1) to get the equivalent uniform annual charge. Figure 13-2 shows graphically the actual and equivalent carrying charges for a capital investment of a piece of equipment with a 5-year life and an assumed 8% cost of money.

The total carrying charges with 8% cost of money for various service lives are estimated as follows:

Years of life	Level annual total carrying charge in %
5	30.82
10	20.59
15	17.44
20	16.04
25	15.34
30	14.96
35	14.76
40	14.67
45	14.63
50	14.63

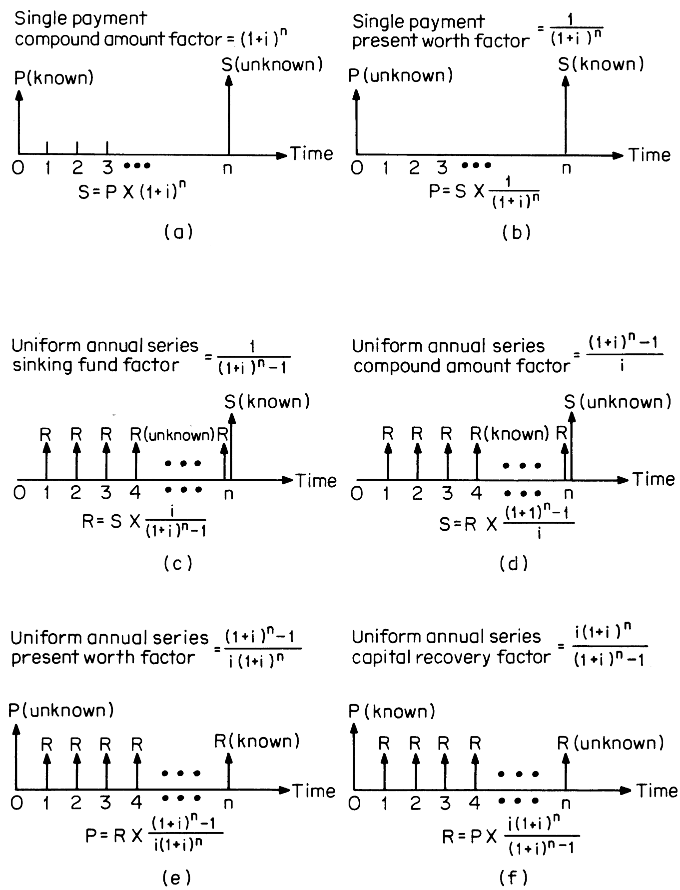


FIGURE 13-1 Graphic interpretations of compound interest factors.

Operating and Maintenance Expenses. This cost component varies with the nature of the project. It is usually not a direct function of the capital invested and may have an inverse tendency. That is, alternatives often exist for higher capital expenditures to reduce operating costs. Therefore, it is *not* expressed as a percent of capital investment in most cases. Nevertheless, it should be included in annual costs.

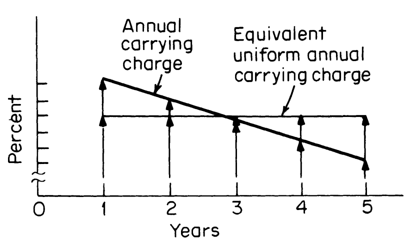


FIGURE 13-2 Representation of carrying charges.

Study Period. When determining the economic comparison of alternatives by comparing the present worth of annual costs, the study period should be taken to the point that the alternatives are equivalent in capability. If this is not practical, the study should be taken so far into the future that the difference in present worth would be insignificant.

Economic Evaluations. A simple example will show a comparison between two alternatives. Let *CC* represent the capital investment multiplied by the level annual carrying charge, operating and maintenance (*O&M*) represent annual operation and maintenance, and *RR* represent the total revenue requirement necessary annually to carry the project. A pad-mounted sectionalizing switch is needed for an underground circuit. The choice is between two manufacturers who can supply the switch but with different characteristics as follows:

	Mfr. A	Mfr. B
Installed cost	\$3600	\$3300
Operating and maintenance	50/year	100/year
Expected life	30 years	20 years

There is no salvage value at end of life. Determine which alternative is less expensive.

The first step is to draw a time diagram like Fig. 13-3. The common point in time for the two alternatives is 60 years, so two cycles of *A* should be compared with three cycles of *B*.

Present-worth analysis:

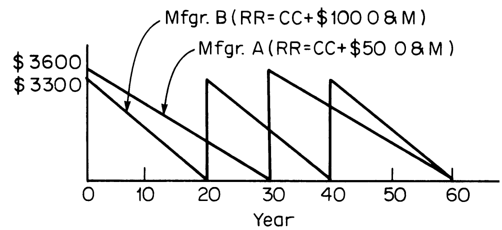


FIGURE 13-3 Time diagram.

$$\begin{aligned}\text{PW Mfr. } A\text{'s alternative} &= 3600 \times 0.1496 \times 11.258 + 50 \times 11.258 \\ &\quad + (3600 \times 0.1496 \times 11.258 + 50 \times 11.258)\, 0.0994 \\ &= 6063.11 + 562.90 + 658.63 \\ &= \underline{\underline{7284.64}}\end{aligned}$$

$$\begin{aligned}\text{PW Mfr. } B\text{'s alternative} &= 3300 \times 0.1604 \times 9.818 + 100 \times 9.818 \\ &\quad + (3300 \times 0.1604 \times 9.818 + 100 \times 9.818)\, 0.2145 \\ &\quad + (3300 \times 0.1604 \times 9.818 + 100 \times 9.818)\, 0.0460 \\ &= 5196.86 + 981.80 + 1325.32 + 284.22 \\ &= \underline{\underline{7788.20}}\end{aligned}$$

where 3600 = installed cost of Mfr. *A*'s switch
0.1496 = level annual carrying charge for 30-year *A* switch
11.258 = 8%, 30-year uniform annual series present-worth factor
50 = *O&M* of *A*'s switch
0.0994 = 8%, 30-year single-payment present-worth factor
3300 = installed cost of Mfr. *B*'s switch
0.1604 = level annual carrying charge for 20-year *B* switch
9.818 = 8%, 20-year uniform annual series present-worth factor
100 = *O&M* of *B*'s switch
0.2145 = 8%, 20-year single payment present-worth factor
0.0460 = 8%, 40-year single payment present-worth factor

Manufacturer *A*'s switch would be the overall lowest cost and would be the better deal provided the capability and reliability of the two switches are equivalent.

13.3 DECISION CRITERIA

There are two measures of the relative worth of projects—the net spendable *amount* of the return (the NPV) and the rate of return on the investment required. The latter measure is the *internal rate of return*. Mathematically, the internal rate of return is the discount rate at which the present value of the cost stream (including both original investments and subsequent costs) equals the present value of the revenue stream. The internal rate of return of the preceding project is obviously greater than 10%, since the NPV is positive at a 10% discount rate. If the NPV had been negative, then it would have been obvious that the internal rate of return was less than 10%.

A decision criterion often used to discriminate between projects is the *payback period,* or *payback*. Mathematically, the payback period is the *cost of the improvement* divided by the *average annual savings*. Although first-year savings are sometimes used as the divisor, the average savings should be used and should include escalations over the life of the project. Using only the first-year savings can yield an incorrect payback.

The following discussion demonstrates that a payback criterion often can lead to the wrong conclusion. If cash stream *A* and cash stream *B* of Table 13-2 were both "savings" streams resulting from the investment of $400 in projects A and B, respectively, the payback would mathematically be the same for each project because they have the same total savings. The average savings (income) is $650 divided by 4 years, or $162.50 per year. The payback for each project is the $400 investment divided by the average annual savings of $162.50, or almost 2.5 years. However, the NPV of each is not equal. The NPV of project A is $100 ($500 to $400); project B's NPV is $130 ($530 to $400). The time value of money causes project B to clearly be the better project; the payback criterion fails to differentiate between the two.

Because of the time-value-of-money problem, a payback criterion actually can indicate that a lesser project is better. For example, if the 1987 savings of project A increased from $220 to $240, the NPV of the project would increase from $100 to $114. Clearly, project B with an NPV of $130 is still better, if the discount rate is 10%. However, the payback period for project A would now decrease from 2.5 to 2.4 years; as a result, the wrong project would be picked if a payback criterion is used.

The type of payback discussed earlier is called a *simple payback* because it uses nominal-year dollars in the calculations. If real (constant-year) dollars are used, it is called the *discounted payback period* or the "breakeven period." In the preceding example, using a discounted payback criterion would have indicated the correct choice in both cases. In Table 13-2, the average discounted savings for projects A and B would be $125 [$500 present value (PV)/4 years] and $132.50, respectively; the discounted paybacks would be 3.2 years ($400/$125/year) and 3 years, respectively. Project B would be chosen because of its shorter payback period.

If the year 4 savings of project A increased to $240, the PV of savings would only increase to $514. Since this would still be less than the PV of $530 for the savings from project B, project A would have lower average discounted savings and a longer discounted payback than project B; the correct relative choice would be made. It is clear that if paybacks are used at all, the discounted payback should be used.

Although the preceding illustration shows the possible folly in looking only at nominal numbers, Table 13-3 and Fig. 13-4 show that folly even better. Both project *X* and project *Y* require a $1000 initial investment. It should be clear from Table 13-3 that project *Y* is the better of the two investments. It would be chosen whether the decision criterion was NPV, internal rate of return, calculated discounted paybacks, or calculated simple paybacks. However, if the first-year savings is used in the payback calculation, or if actual payback time (see the graph) is used, project *X* would be chosen. This shows the problem with using first-year savings instead of average savings; it also brings up another important point. It is *cash flows* which dominate business decisions; both the level and the timing of those flows can be critical. Project *X* could very well be the appropriate project to choose if the timing of its cash flows allowed other projects to be undertaken such that the aggregate benefit of all projects was increased. The final decisions on projects should be made on an overall benefit basis.

Another useful tool for comparing projects is the *benefit-cost ratio,* which is the present value of the benefits (savings) divided by the initial cost. For projects *X* and *Y* of Table 13-3, the benefit-cost

TABLE 13-3 Net Present Value vs. Payback ($1000 original investment, 10% discount rate)

						Project return				
	Net cash returns by year					Nominal $		Discounted $		
Project	1	2	3	4	5	Total	Net	PV	NPV	%IRR*
X	500	500	250	100	50	1400	400	1155	155	18.5
Y	200	300	400	500	600	2000	1000	1444	444	23.3

	Simple payback, year			Discounted paybacks, year	
	Calculated using				
Project	1st-year savings	Average savings	Actual payback	Calculated	Actual
X	2.0	3.6	2.0	4.3	3.5
Y	5.0	2.5	3.2	3.5	3.8

*IRR = internal rate of return.

ratios are 1.155 and 1.444, respectively. When the appropriate discount rate is used, any benefit-cost ratio greater than unity (1.0) indicates that the project is profitable.

Calculating the NPV and the internal rate of return from each alternative project is a rational method of discriminating between projects and ranking them in an investment priority. First, the projects can be ranked in descending order by the internal rates of return. With an unlimited amount of money and management time, a company would be expected to start all projects with an internal rate of return greater than the cost of money to the company. However, in the "real world," this is usually not the case. The firm is usually limited in capital, or in management capability, and must choose a

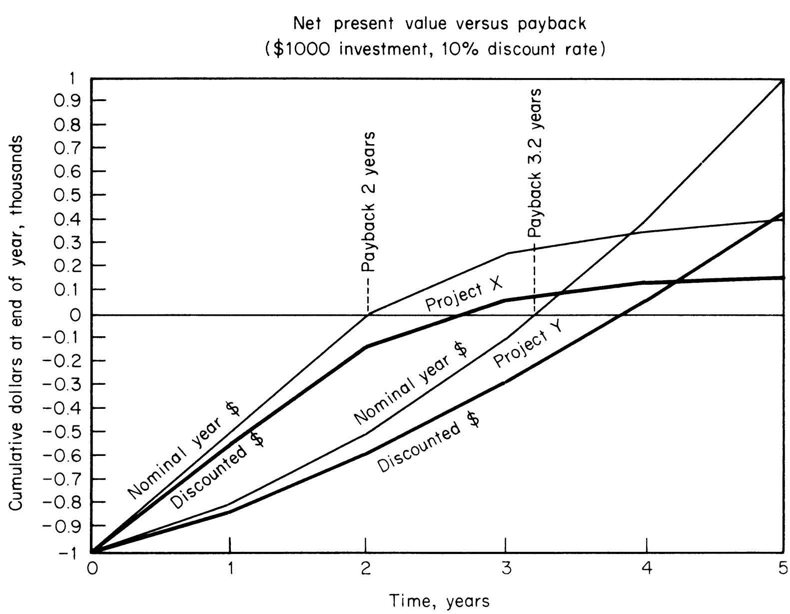


FIGURE 13-4 Graph of net present value vs. payback (values from Table 13-3).

subset of the complete menu of alternative projects. The NPVs of the projects can be used to help match available resources to achieve the greatest total real return.

In addition to the consideration of the real income and the real rates of return from the various projects, the nominal-dollar flows of each project must be considered to ensure that the cash flow of the company will be great enough to provide the capital needed in each time period.

If the total cash outlay required for all projects is greater than the total income during any period, the company must either borrow the shortfall or pay it out of available cash. For many companies, available cash is tight, and expected business conditions are not good or are uncertain. These companies will rarely invest in a set of projects that may put them in financial jeopardy—even if the expected long-term returns are great. It is not unusual for a low-return project to be substituted for a high-return project when the cash requirements of the high-return project coincide with other cash demands and the company cannot economically provide the required funds at that time.

The example in Table 13-3 is simplistic. It incorrectly assumes that (1) the project costs and returns are certain and (2) all proceeds of the project can be retained by the owner. Uncertainty of cash flows should be considered by using "sensitivity analysis" and comparing expected results under both optimistic and pessimistic conditions. The tax consequences of the manner in which a project is financed are discussed in the next sections. Further comments on the characteristics affecting the type and amount of an investment are provided at the end of this primer.

13.4 AFTER-TAX CASH FLOWS

The net amount of cash available for reinvestment in the company or distribution to the owners depends on the tax consequences of a project and its financing. For tax purposes, there are two kinds of project expenditures—expensed and capitalized. Expenditures for short-lived items consumed in making a product or providing a service are generally allowed to be "expensed" in the year they are made. Such expenses are allowed to be deducted from gross income before taxes are computed. Examples are rent, parts, travel expenses, utility bills, raw materials, labor, and advertising.

Capitalized expenditures will continue to give service for several years. The company is allowed to recover those expenses over a number of years by deducting a percentage of the cost each year from the gross income of the company before calculating the taxes. This "depreciation recovery" follows specific rules for the number of years over which the recovery is made and the percentage of the cost allowed as a tax deduction each year.

Typical capitalized expenditures are buildings, machinery, and land. Since buildings and machinery are "consumed" in service, they are considered depreciable property. Land, however, is not consumed and cannot be depreciated except under special circumstances, such as where the usefulness of the land is indeed consumed and a depletion allowance is authorized.

Deductions have no value in themselves; they merely serve to reduce the amount of income that is taxable. As a result, the actual value of an allowed expense or depreciation deduction depends on the *incremental tax rate* of the company. This is the rate charged against the "last" income earned in a year. Since a tax deduction offsets or "shelters" income by reducing the taxable income, the value of a tax deduction is the amount of tax that would have been paid on the income that is sheltered by the deduction. The higher the incremental tax rate, the greater the tax expense avoided by taking the deduction. The reduction in income taxes that *results* from allowed deductions has the same effect as an increase in project revenues; each increases the net revenues of the project. (*Note:* Deductions are not cash items and are not spendable income; their value is that they generate savings in taxes that otherwise would have to be paid.)

Most projects qualify for one or more special tax subsidies called *tax credits*. A tax credit can offset a tax otherwise owed to the government; the actual cash required for paying taxes is thus reduced. Tax credits are usually in the form of a stated percentage of the capitalized project investment and are usually allowed only in the year of the investment. Unlike allowed depreciation, the effect on the company from a tax credit is independent of the incremental tax rate. The tax credit is a direct reduction in the tax liability of the company. If the tax credit is greater than the tax liability in that year, the unused portion can be applied in other years.

TABLE 13-4 Tax Calculation

Gross income
− interest payments
− operating expenses
− allowable amortization and depreciation on equipment
− other tax-deductible expenses
= taxable income (+ or −)
× incremental tax rate
= initial tax liability (+ indicates due, − indicates saved)
− total tax credits (only if tax liability is positive)
= actual tax (+ indicates due, − indicates saved)

Note: Taxable income is the net difference between gross income and allowed deductions. Since taxable income determines the actual tax liability, it is easy to see the effect on after-tax income of increasing or decreasing the allowed deductions.

The net cash flow in spendable dollars yielded by a project depends on the gross income and the cash expenditures which must be made as a result of the project. The tax effects of the investment and the method of financing the investment can sometimes "make or break" a project. Table 13-4 shows the items that must be considered when calculating tax liabilities.

Table 13-5 shows two methods of calculating the effect of taxes on cash flow; both yield the same answer. These methods are presented here to aid in understanding the effect of nondeductible expenses and noncash tax deductions on the cash flow of a given year. Principal payments on loans are not allowed as a tax deduction, but they are cash payments that must be made during the year. On the other hand, depreciation on depreciable assets is allowed as a tax deduction, and therefore reduces taxes, but it is not an out-of-pocket cash expenditure.

TABLE 13-5 Cash Flow Calculations

Method 1
Taxable income
− principal payments on debt
+ allowable amortization and depreciation (these are noncash-deductible expenses and, as such, are not spent but available)
= cash available for taxes
− tax due (or + tax savings)
= after-tax cash income
Method 2
Gross income
− interest payments
− principal payments
− other cash expenses
= cash available for taxes
− tax due (or + tax savings)
= after-tax cash income

Note: These calculations assume that the total income of this project and other projects is great enough for the owner to use all of the benefits earned in this year. Otherwise, some of the benefits may be carried into another tax year—but they will be worth less because of the time value of money.

13.5 FINANCING EFFECTS

The examples in Table 13-6 show the tax benefits that result from changing the method of financing a project. The project requires an initial investment of $4000. If as in line 1 the owner finances the whole project with personal equity funds, without borrowing any funds and going into debt, the only tax deduction allowed over the life of the project is the depreciation expense. Since both the tax credits and the depreciation expense are related only to the cost of the depreciable assets, and not to the method of financing, both are the same in all cases. If the owner has a 50% incremental tax rate, the allowed deductions generate $2000 in tax savings if the project is 100% equity-financed. The resulting tax benefits total 60% of the original equity investment.

If the owner borrows $1000 and invests $3000 of his or her own money, that is, finances the project in a 25:75 debt-equity ratio, the allowed tax deductions rise by the $492 interest deduction, and the tax benefits increase. Financing part of a project with debt funds is called *leveraging* the equity investment. All the benefits of the project continue to flow to the owner, and the tax benefits themselves increase. As a result of the increased benefits and the decreased equity investment, the ratio of tax benefits to equity increases; the rate of return on the investment thus increases, even though the project itself is bringing in the same gross income.

If the project is financed with a 75:25 debt-equity ratio, the tax benefits which accrue to the owner amount to over 3 times its original equity investment. There are no free lunches, however. If the project fails to reach its income objectives, or costs run higher than expected, the owner will still be liable for payment of the principal and interest payments on the money borrowed for the project. The higher the leverage of the investment in the project, the higher is the business risk the owner faces.

Table 13-7 contains the data for the illustrations of financing effects in the remaining tables. The payments for principal and interest are shown for a debt of $1000 to be repaid over 5 years at 15% interest. The depreciation rates allowed under the accelerated cost recovery system (ACRS) are shown along with the annual depreciation and the investment tax credit allowed on a $2000 depreciable investment. The tables in this text were prepared using 1982 regulations and have been retained for simplicity of illustration. Since tax credits and tax deductions change frequently, care should be taken to use the correct allowances.

Table 13-8 shows the calculations of tax effects and cash flows for a $2000 project which the owner finances completely with equity investment. There are no interest deductions included in the tax calculations, since there is no debt to repay. Likewise, there are no principal payments included in the cash flow calculations. The incremental income tax rate of the owner is assumed to be 50%. This method of financing the project yields a nominal return of $4434 over 5 years from an original investment of $2000. The internal rate of return is 32.6%.

Table 13-9 shows the same project, except that it is now financed with 50% equity and 50% debt, with the debt cost assumed at a rate of 15% per year. A 50:50 debt-equity ratio increases the cash outflow required to service the debt; it reduces the overall nominal return over the 5 years to $3187. However, since the owner invested only $1000, the internal rate of return of the project increases to the 55% level. This indicates that if the owner had $2000 to invest, it would be better (other things

TABLE 13-6 Examples of Tax Benefits
Total tax benefits received by owner of a $4000 project

Percent equity financing	Owner equity investment, $	Amount borrowed, $	Depreciation expense deduction, $	Interest expense deduction $ 15%, $	Total deductions, $	Taxes saved @50%, $	Inv. tax credits, $	Total cash benefits, $	Ratio tax benefits-equity
100	4000	0	4000	0	4000	2000	400	2400	0.60
75	3000	1000	4000	492	4492	2246	400	2646	0.88
50	2000	2000	4000	983	4984	2492	400	2892	1.45
25	1000	3000	4000	1476	5476	2738	400	3138	3.14

TABLE 13-7 Data for Illustrations in Tables 13-8 through 13-14
Payments based on $1000 borrowed for 5 years at 15% interest

	Annual cash payments by year, $					
	1	2	3	4	5	Total
Interest (tax-deductible)	150.00	127.75	102.17	72.75	38.91	491.58
Principal (not deductible)	148.32	170.57	196.15	225.57	259.41	1,000.02
Payment	298.32	298.32	298.32	298.32	298.32	1,491.60
	Year					
	1	2	3	4	5	
ACRS depreciation rates, %	15	22	21	21	21	
Allowed depreciation deduction on $2000	300	440	420	420	420	
10% investment tax *credit* (not deduction)						
First year only	200					

Note: Assumed combined federal and state tax rate = 50%

TABLE 13-8 100% Owner Financing of a $2000 Project

	Year					
	1	2	3	4	5	Total
Tax calculations						
Revenues	1500	1680	1880	2100	2360	9520
− interest	0	0	0	0	0	0
− O&M expenses*	500	550	605	666	732	3053
− depreciation	300	440	420	420	420	2000
= taxable income	700	690	855	1014	1208	4467
× tax rate	0.50	0.50	0.50	0.50	0.50	0.50
= initial tax due	350	345	427	507	604	2233
− tax credits	200	0	0	0	0	200
= actual tax due	150	345	427	507	604	2033
Cash flow						
Taxable income	700	690	855	1014	1208	4467
− principal payments	0	0	0	0	0	0
= depreciation	300	440	420	420	420	2000
= cash available	1000	1130	1275	1434	1628	6467
+ tax due	150	345	427	507	604	2033
= after-tax cash income	850	785	848	927	1024	4434

Note: The original owner investment of $2000 returns over $4000 in 5 years for an internal rate of return of 32.6%.
*Operation and maintenance.

TABLE 13-9 50% Debt and 50% Owner Financing of a $2000 Project

	Year					
	1	2	3	4	5	Total
Tax calculations						
Revenues	1500	1680	1880	2100	2360	9520
− interest	150	128	102	73	39	492
− O&M expenses	500	550	605	666	732	3053
− depreciation	300	440	420	420	420	2000
= taxable income	550	562	753	941	1169	3975
× tax rate	0.50	0.50	0.50	0.50	0.50	0.50
= initial tax due	275	281	376	471	585	1988
− tax credits	200	0	0	0	0	200
= actual tax due	75	281	376	471	585	1788
Cash flow						
Taxable income	550	562	753	941	1169	3975
− principal payments	148	171	196	226	359	1000
+ depreciation	300	440	420	420	420	2000
= cash available	702	831	977	1135	1330	4975
− tax due	75	281	376	471	585	1788
= after-tax cash income	627	550	601	664	745	3187

Note: The original *owner* investment of $1000 returns over $3000 in 5 years for an *internal rate of return* of 54.5%. *Leveraging* the owner's equity investment 1:1 with debt causes the internal rate of return on the owner's equity investment to rise because the owner invests only half the money but still receives the full tax benefits.

being equal) to invest $1000 each in two such projects. The yield would then be $6374 for a $2000 investment, as compared with $4434 if only one project is completely owner-financed.

Table 13-10 shows that the same project, with a 30% owner tax rate, yields $3982 in income for the $1000 initial investment. (*Note:* All the tax credit could not be used in the first year because the tax liability was reduced by the lower tax rate.)

TABLE 13-10 50% Debt and 50% Owner Financing of a $2000 Project, with an Owner Tax Rate of 30%

	Year					
	1	2	3	4	5	Total
Tax calculations						
Taxable income	550	562	753	941	1169	3975
× tax rate	0.30	0.30	0.30	0.30	0.30	0.30
= initial tax due	165	169	226	282	351	1193
− tax credits	165	35	0	0	0	200
= actual tax due	0	134	226	282	351	993
Cash flow						
Cash available	702	831	977	1135	1330	4975
− tax due	0	35	226	282	351	993
= after-tax cash income	702	697	751	853	979	3982

13.6 LEASING

When a lease arrangement is worked out between two parties, the lessor party owns the installation, and the lessee party pays for its use. Since the lessee must pay enough profit to the lessor for the lessor to be willing to install the property for the lessee's use, this arrangement might not appear advantageous to the lessee. However, leasing can be a great advantage in several situations, particularly when the lessee does not want to or cannot borrow the initial money required. The tax advantages of a lease often make a project *go* with lease financing when it cannot go otherwise. Tables 13-11 to 13-14 examine the cash flows that occur in a leasing situation.

Table 13-11 calculates the revenue required for the lessor to recover its expenses and investment *without* any return, that is, to break even, if it installs the project and leases it to a lessee. In this particular case, the lessor would make no profit and there would be no incentive to install the project. This case is shown only for the purpose of having a clean example to use as a base for leading into the following examples. Table 13-11 shows the effect of the tax deductions on the lessor; it also shows the out-of-pocket expenses of the lessor that must be covered by the lessee if the lessor breaks even. This is essentially the same set of calculations shown in Table 13-9, except that Table 13-11 calculates the breakeven point.

If the lessor breaks even, Table 13-12 shows the return to the lessee from leasing the project from the lessor. In this case, the lessee invests no money in the project and still reaps a handsome reward. One of the mechanisms that makes leasing work is that the lessee can take the entire cost of the lease as a deduction before taxes, *including the cost of the principal payments of the lessor*. If the lessee were to put the project in on its own, as in Table 13-9, it could deduct only depreciation and interest payments. By leasing, the lessee gets, in effect, two bites at the apple; it gets to deduct the entire

TABLE 13-11 Required Breakeven Revenue for Lessor for \$2000 Project with 50:50 Debt-Equity Ratio

	Year					
	1	2	3	4	5	Total
Tax calculations						
Interest	150	128	102	73	39	492
+ O&M expenses	500	550	605	666	732	3053
+ depreciation	300	440	420	420	420	2000
= deductible expenses	950	1118	1127	1159	1191	5545
× tax rate	0.50	0.50	0.50	0.50	0.50	0.50
= initial taxes saved	475	559	563	580	595	2772
+ tax credits	200	0	0	0	0	200
= actual taxes saved	675	559	563	580	595	2972
Cash flow						
Interest & principal	298	299	298	299	298	1492
+ O&M expenses	500	550	605	666	732	3053
− tax savings	675	559	563	580	595	2972
= operating cash outlay	123	290	340	385	435	1573
+ investment recov	200	200	200	200	200	200
= required cash	323	490	540	585	635	2573
× 2 (tax factor)						
= required revenue	646	980	1080	1170	1270	5146

TABLE 13-12 Income of Lessee if Lessor Breaks Even

	Year					
	1	2	3	4	5	Total
Tax calculations						
Revenues	1500	1680	1880	2100	2360	9520
− lease payment	646	980	1080	1170	1270	5146
= taxable income	854	700	800	930	1090	4374
× tax rate	0.50	0.50	0.50	0.50	0.50	0.50
= tax due	427	350	400	465	545	2187
Cash flow						
Revenues	1500	1680	1880	2100	2360	9520
− income taxes	427	350	400	465	545	2187
− lease payment	646	980	1080	1170	1270	5146
= after-tax cash income	427	350	400	465	545	2187

Note: The long-run economics of leasing would depend upon the terms of the lease and the residual ownership and use of equipment after initial payoff.

TABLE 13-13 Required Lessor Revenue if Lessor Makes 15% on Investment on a $2000 Project with 50:50 Debt-Equity Ratio

	Year					
	1	2	3	4	5	Total
Tax calculations						
Interest	150	128	102	73	39	492
+ O&M expenses	500	550	605	666	732	3053
+ depreciation	300	440	420	420	420	2000
= deductible expenses	950	1118	1127	1159	1191	5545
× tax rate	0.50	0.50	0.50	0.50	0.50	0.50
= initial taxes saved	475	559	563	580	595	2772
+ tax credits	200	0	0	0	0	200
= actual taxes saved	675	559	563	580	595	2972
Cash flow						
Loan payment ($i + p$)	298	299	298	299	298	1492
+ O&M expenses	500	500	605	666	732	3053
− tax savings	675	559	563	580	595	2972
= operating cash outlay	123	290	340	385	435	1573
+ recovery of initial investment @ 15% return	298	299	298	299	298	1492
= required cash	421	589	638	684	733	3065
× 2 (the tax factor)						
= required income	842	1178	1276	1368	1466	6130

TABLE 13-14 Income of Lessee if Lessor Makes 15%

	Year					
	1	2	3	4	5	Total
Tax calculations						
Revenues	1500	1680	1880	2100	2360	9520
− lease payment	842	1178	1276	1368	1466	6130
= taxable income	658	502	604	732	894	3390
× tax rate	0.50	0.50	0.50	0.50	0.50	0.50
= tax due	329	251	302	366	447	1695
Cash flow						
Revenues	1500	1680	1880	2100	2360	9520
− income taxes	329	251	302	366	447	1695
− lease payment	842	1178	1276	1368	1466	6130
= after-tax cash income	329	251	302	366	447	1695

Note: With a zero investment by the lessee, the lessor makes a 15% return and the lessee still makes $1695, an infinite return. The long-run economics of leasing would depend upon the terms of the lease and the residual ownership and use of equipment after the initial payoff.

lease payment before taxes. Since the lease payment includes both the principal payments and the tax effects of depreciation allowances, the lessee, in effect, gets to write off the project twice, once at the lessor's incremental tax rate and once at the lessee's incremental tax rate.

Table 13-13 is the same as Table 13-11, except that Table 13-13 calculates the revenue required to produce a 15% return on investment for the lessor, rather than a breakeven return. Required income almost doubles, primarily because of the income taxes that have to be paid on taxable income before the net cash is available to the lessor.

Table 13-14 shows that the effect of allowing the lessor to earn a 15% rate of return is to cut the lessee's after-tax income roughly in half. However, since the lessee still hasn't invested any money in the project, the rate of return of the lessee is infinitely large. When the return of $1695 from leasing is compared with the return of Table 13-9, where an initial investment of $1000 is required, the attractiveness of many leasing schemes is immediately seen. When such schemes are combined with provisions for the lessee to be able to buy the project from the lessor in the future at a reasonable price and at lessee's option, the package can be especially attractive.

In some cases, leasing is used to protect the lessee from buying a set of equipment that may not work well for its application. By leasing, the lessee gets a chance to work with the equipment and see if it performs as expected—before spending large amounts of investment capital on the installation.

13.7 RATE-OF-RETURN REQUIREMENTS

There are three components of *interest rates*. The first is the *liquidity factor*. There is a value in having cash available to use for whatever investment opportunity may appear in the future. Before one person will lend money to another, the interest earned must compensate the lender for the unavailability of its money while the borrower still has it, that is, for the lack of liquidity. Second, just like one neighbor lending another a lawn mower, the lender of money expects to get it back in just as valuable a condition as when it was borrowed. In the case of money, the borrower must increase the interest rate paid to the lender enough to include the expected rate of *inflation*. This allows the lender to recover the same *value* as originally lent, albeit a greater number of dollars. The third item that must be included in the interest rate, before a lender is willing to part with the money, is enough additional interest to offset any *risks* associated with the loan. Obviously, the riskier a loan appears, the higher the interest rate required by the lender will be. All of these factors entail *uncertainty*. The lender is uncertain about what opportunities may come along later, the devaluation of the loan from inflation, the ability of the lender to repay the loan, changes in government regulations, and other factors.

These same factors influence the *minimum expected rate of return,* or the *hurdle rate,* that a company requires a project to meet or exceed before giving it full consideration. If the company must borrow money to finance the project, it will be concerned about its ability to repay the loan without jeopardizing the company. The very financing methods which leverage a company's investment and produce such high possible returns also leverage the company's financial risk. Usually, the more stable the expected earnings from projects, the more leverage the company is willing to risk.

If a hurdle rate is used to screen potential projects, the hurdle rate should appropriately reflect the weighted cost of capital to the firm. Using a hurdle rate that is significantly different from the weighted cost of capital incorrectly rejects and accepts projects.

It is not correct to use either the opportunity cost of using retained earnings or the interest rate on borrowed debt solely as the hurdle rate. If retained earnings are used in one project, they are unavailable for use in others. The opportunity cost of using those funds is the rate of return that could be earned by investing those funds in routine company business opportunities. As such, they are generally both higher cost and less extensive than available debt funds. Using that rate can deny worthwhile projects and choke the expansion of the firm.

Considering a project to be financed entirely by debt is also unrealistic. If funds are borrowed without a complementary equity investment, the debt-equity ratio rises, the debt coverage ratio falls, and the ability to borrow more funds decreases.

As a result of the above and related factors, the appropriate hurdle rate is the weighted cost of capital to the firm.

Hurdle rates are often used both as a threshold of profitability that projects must meet and as a method of discriminating *between* projects. As stated earlier, using a hurdle rate that is significantly different from the actual cost of capital to the firm will undercommit or overcommit the firm. As shown below, it may also lead to an incorrect choice of projects.

If all projects under consideration have positive cash flows in later years, almost any hurdle rate can be used to determine the "best" projects on a relative NPV basis. The higher the hurdle rate used to discount future cash flows, the lower the resulting NPV. The result may be the wrong NPV, but the relative ranking will not change. However, that is *not* the case where one or more of the projects have some later years with negative cash flows, such as when significant investments in maintenance or replacement are required; relative rankings may change.

In effect, the discount rate used as a hurdle is assumed to be a rate that can continue to be earned in other areas by the dollars returned from a project each year. It can be used to pay off debt and "earn" the avoided interest, or it can be put into another income-producing project. In addition, the higher the discount rate, the less value are later revenues. It is these effects which require the hurdle rate to be close to the actual cost of capital. If a firm cannot actually earn the hurdle rate by reinvesting each year's proceeds from a project, the wrong project may be chosen.

13.8 CHARACTERISTICS AFFECTING INVESTMENTS

The preceding discussions have briefly covered some of the factors that drive the decisions people make about new projects and affect the amounts and types of investments. The following is a summary of items that must be considered when any major project is examined:

Ability to borrow money

Cash on hand

Relative risk of the project

Ability to use tax benefits

Existence of tax credits or unusual benefits or constraints

Relative tax rates

Timing of the costs and revenues

Relative permanence of the investment

Ability to shift to another investment if one becomes more attractive

Ability to maintain and operate the project equipment

The ability of a party with money and a party with a project need to find a satisfactory arrangement for (1) financing the project and (2) appropriately sharing the risks and the benefits is almost limitless. Both parties (they may be the same party if the project is primarily owner-financed) must find an acceptable level of risk and reward.

13.9 RISK AND REWARD

To many people, taking a risk is its own reward; to others, very little risk is worth taking. The successful manager will analyze alternative projects and will adjust project parameters and financing methods to yield combinations of risk and expected reward appropriate for all parties.

The successful project analyst will be guided by the TANSTAAFL principle: *There ain't no such thing as a free lunch*. Someone, somewhere pays for everything. The questions are who? how much? and when? Answering these provides the basis for sound decisions.

BIBLIOGRAPHY

Caywood, R. E.: *Electric Utility Rate Economics.* New York, McGraw-Hill, 1972.

Childs, J. F.: *Encyclopedia of Long-Term Financing and Capital Management.* Englewood Cliffs, N.J., Prentice-Hall, 1976.

Clapp, A. L.: *Primer on Project Economics.* Research Triangle Park, N.C., North Carolina Alternative Energy Corporation, 1984.

Schall, L. D., and Haley, C. W.: *Introduction to Financial Management.* New York, McGraw-Hill, 1977.

Weston, J. F., and Brigham, E. F.: *Managerial Finance,* 4th ed. Hinsdale, Ill., Aryden Press, 1972.

SECTION 14
TRANSMISSION SYSTEMS

E. C. (Rusty) Bascom, III
Senior Engineer, Power Delivery Consultants, Inc.; Senior Member, IEEE

J. R. Daconti
Executive Consultant, Siemens Power Technologies International; Senior Member, IEEE, Distinguished Member, CIGRE

D. A. Douglass
Principal Engineer, Power Delivery Consultants, Inc., Fellow, IEEE

A. M. DiGioia, Jr.
Chairman Emeritus, GAI Consultants, Inc.; Fellow, ASCE; Member, IEEE

I. S. Grant
Manager, TVA; Fellow, IEEE

J. D. Mozer
Staff Consultant, GAI Consultants, Inc.; Member ASCE

J. R. Stewart
Consultant; Fellow IEEE

J. A. Williams
Principal Engineer, Power Delivery Consultants, Inc.; Fellow, IEEE

CONTENTS

14.1 OVERHEAD AC POWER TRANSMISSION

Overhead transmission of electric power remains one of the most important elements of today's electric power system. Transmission systems deliver power from generating plants to industrial sites and to substations from which distribution systems supply residential and commercial service. Those transmission systems also interconnect electric utilities, permitting power exchange when it is of economic advantage and to assist one another when generating plants are out of service because of damage or routine repairs. Total investment in transmission and substations is approximately 10% of the investment in generation.

Since the beginning of the electrical industry, research has been directed toward higher and higher voltages for transmission. As systems have grown, higher-voltage systems have rarely displaced existing systems, but have instead overlayed them. Economics have typically dictated that an overlay voltage should be between 2 and 3 times the voltage of the system it is reinforcing. Thus, it is common to see, for example, one system using lines rated 115, 230, and 500 kilovolts (kV). The highest ac voltage in commercial use is 765 kV although 1100 kV lines have seen limited use in Japan and Russia. Research and test lines have explored voltages as high as 1500 kV, but it is unlikely that, in the foreseeable future, use will be made of voltages higher than those already in service. This plateau in growth is due to a corresponding plateau in the size of generators and power plants, more homogeneity in the geographic pattern of power plants and loads, and adverse public reaction to overhead lines. Recognizing this plateau, some focus has been placed on making intermediate voltage lines more compact. Important advances in design of transmission structures as well as in the components used in line construction, particularly insulators, were made during the mid-1980s to mid-1990s. Current research promises some further improvements in lines of existing voltage including uprating and now designs for HVDC.

14.1.1 Transmission Systems

The fundamental purpose of the electric utility transmission system is to transmit power from generating units to the distribution system that ultimately supplies the loads. This objective is served by transmission lines that connect the generators into the transmission network, interconnect various areas of the transmission network, interconnect one electric utility with another, or deliver the

TABLE 14-1 Standard System Voltages, kV

Rating		Rating	
Nominal	Maximum	Nominal	Maximum
34.5	36.5	230	242
46	48.3	345	362
69	72.5	500	550
115	121	765	800
138	145	1100	1200
161	169		

electrical power from various areas within the transmission network to the distribution substations. Transmission system design is the selection of the necessary lines and equipment which will deliver the required power and quality of service for the lowest overall average cost over the service life. The system must also be capable of expansion with minimum changes to existing facilities.

Electrical design of ac systems involves (1) power flow requirements; (2) system stability and dynamic performance; (3) selection of voltage level; (4) voltage and reactive power flow control; (5) conductor selection; (6) losses; (7) corona-related performance (radio, audible, and television noise); (8) electromagnetic field effects; (9) insulation and overvoltage design; (10) switching arrangements; (11) circuit-breaker duties; and (12) protective relaying.

Mechanical design includes (1) sag and tension calculations; (2) conductor composition; (3) conductor spacing (minimum spacing to be determined under electrical design); (4) types of insulators; and (5) selection of conductor hardware.

Structural design includes (1) selection of the type of structures to be used; (2) mechanical loading calculations; (3) foundations; and (4) guys and anchors.

Miscellaneous features of transmission-line design are (1) line location; (2) acquisition of right-of-way; (3) profiling; (4) locating structures; (5) inductive coordination (considers line location and electrical calculations); (6) means of communication; and (7) seismic factors.

14.1.2 Voltage Levels

Standard transmission voltages are established in the United States by the American National Standards Institute (ANSI). There is no clear delineation between distribution, subtransmission, and transmission voltage levels. In some systems 69 kV may be a transmission voltage while in other systems it is classified as distribution, depending on function. Table 14-1 shows the standard voltages listed in ANSI Standards C84 and C92.2, all of which are in use at present.

The nominal system voltages of 345, 500, and 765 kV from Table 14-1 are classified as extrahigh voltages (EHV). They are used extensively in the United States and in certain other parts of the world. In addition, 400-kV EHV transmission is used, principally in Europe. EHV is used for the transmission of large blocks of power and for longer distances than would be economically feasible at the lower voltages. EHV may be used also for interconnections between systems or superimposed on large power-system networks to transfer large blocks of power from one area to another.

One voltage level above 800 kV, namely, 1100 kV nominal (1200 kV maximum), is presently standardized. This level is not widely, although sufficient research and development have been completed to prove technical practicability.[*1–3]

14.1.3 Conductor Selection

Considerations in Selection.[4] The choice of a conductor for a transmission line, as with structure type, depends on the specific application. Once the mechanical strength requirement of the conductor

[*]Superscript numbers refer to references listed at the end of this section (*1–3).

is satisfied, the conductor choice considers the total costs associated with the conductor and also the corona-related electrical environmental effects of radio and audible noise. Corona also causes power loss, particularly during wet weather.

The electrical stress on the surface of a conductor is a function of the voltage on the conductor, the size (i.e., surface area) of a conductor, and the spacing between conductors and/or grounded objects.

The equivalent size of a conductor can be increased by using either a larger conductor or several smaller conductors electrically and physically connected together (bundled conductors). While a single, very large conductor would be electrically adequate, several smaller conductors offer practicality of manufacturing and transporting, ease of construction, and minimizing material usage and mechanical stresses on the supporting structures during high winds and/or ice on the conductors.

At voltages of 345 kV and above, the minimum conductor size or the minimum number of conductors and the individual conductor size in a bundle are, in addition to cost considerations, normally determined by the corona-related electrical environmental effects. At voltages below 345 kV (e.g., 69 through 230 kV), the minimum size is normally based only on conductor economics.

The conductor sag in the span between structures will depend on conductor materials, conductor weight, conductor strength, conductor tension, conductor temperature, and ice accumulation on the conductor. Strong conductors can be installed at higher tensions and will sag less.

As the current in a conductor increases, the losses increase with a resultant increase in conductor temperature, causing the sag to increase. If the conductor is carrying heavy electrical load on a hot day, very significant increases in sag can occur. Short spans of 150 to 300 ft may have sags of 2 to 5 ft. Long spans of 1000 to 1500 ft may experience sags of 40 ft or more.

Since a limiting design criterion is minimum conductor height above ground (for safety reasons), the maximum sags during operation can determine structure heights and span lengths. Similarly, in certain areas ice can form on the conductors of sufficient weight to limit the structure heights and span lengths to maintain ground clearance.

Economics. Conductor economic analyses normally use the present worth of revenue required (PWRR) method. This considers the sum of the present worth of levelized annual fixed charges on the total line capital investment, plus annual expenses for line losses:

$$\text{PWRR} = \sum_{n=1}^{\text{NYE}} \left(1 + \frac{i}{100}\right)^{-n} \times \left(\text{CI} \times \frac{F_L}{100} + \text{ADC}_n + \text{AEC}_n\right) \tag{14-1}$$

where PWRR = present worth of revenue required
NYE = number of years to be studied
n = nth year
i = annual discount rate in percent
CI = total per mile capital investment
F_L = line fixed-charge rate in percent
ADC_n = per mile demand charge for line losses for year n
AEC_n = per mile energy charge for line losses for year n

The cost of line losses is based on the cost of generating the losses. Annual demand and energy charges are calculated as shown in the following equations.

Annual demand charge for line losses for year *n:*

$$\text{ADC}_n = \frac{C_{\text{kW}} \times \text{ESC}_n}{10^3} \times \frac{F_g}{100} \times \left[1 + \frac{\text{RES}}{100} \times I_L^2 \times \frac{R}{N_c} \times N_{\text{ckt}} \times N_p\right] \tag{14-2}$$

where ADC_n = annual demand charge for year n
C_{kW} = installed generation cost in dollars per kilowatt
ESC_n = escalation cost factor for year n
F_g = generation fixed-charge rate in percent
RES = required generation reserve in percent

I_L = demand phase current in amperes per circuit
R = single conductor resistance in ohms per mile
N_c = number of conductors per phase
N_{ckt} = number of circuits
N_p = number of phases

Annual energy charge for line losses for year *n:*

$$\text{AEC}_n = \frac{C_{\text{MWh}} \times \text{ESC}_n}{10^6} \times 8760 \times \frac{L_f}{100} \times I_L^2 \times \frac{R}{N_c} \times N_{\text{ckt}} \times N_p \tag{14-3}$$

where AEC_n = annual energy charges for year n
C_{MWh} = cost of generating energy in dollars per megawatthour
ESC_n = escalation cost factor for year n
L_f = loss factor for determining energy losses in percent
I_L = demand phase current in amperes per circuit
R = single conductor resistance in ohms per mile
N_c = number of conductors per phase
N_{ckt} = number of circuits
N_p = number of phases

As the conductor size increases, the installed cost increases, because of both the increased conductor cost and the stronger structures necessary to support the larger, heavier conductor and the attendant mechanical loading. The larger conductor cross section, however, results in lower resistance and therefore lower losses. If corona losses are considered, these are also reduced for larger conductors, assuming other dimensions (e.g., phase spacing) remain constant. Therefore, there will be an overall minimum cost at a specific conductor size, where installed cost forces the PWRR higher for large conductors and the cost of losses forces the PWRR higher for smaller conductors. This is conceptualized in Fig. 14-1. In most practical analyses, there is a relatively flat "minimum" total cost (PWRR) region such that the line designer can temper the economic choice with other factors. Various conductor designs and configurations, such as number of conductors per bundle and size of conductors in a bundle, are examples of areas of designer preference. The higher cost of energy, primarily due to increased fuel costs, has increased the significance of cost of losses in the economic analysis, skewing the economics toward larger conductors with lower losses.

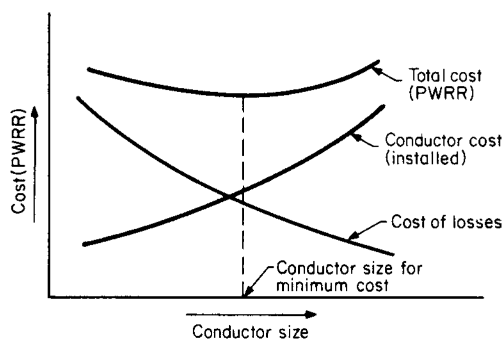


FIGURE 14-1 Conductor economic concept.

Beside the cost of electrical losses, the choice of conductor is an important factor in determining the maximum allowable power flow through the line. For long lines, maximum allowable power flow may be determined by limits on electrical phase shift or voltage drop. For shorter lines, the maximum conductor temperature (thermal rating) may limit the maximum allowable power flow. High thermal capacity can be accomplished either by using a large diameter conductor with relatively low electrical resistance or by using a conductor of relatively smaller diameter tolerant to high operating temperatures, such as ACSS conductors, which can operate continuously at 200°C with no changes in their mechanical properties. For example, consider the following thermal ratings calculated for a transmission line located in an environment with air temperature of 40°C, full sun, and a perpendicular wind speed of 2 ft/s.

Conductor name	Aluminum area (kcmil)	Maximum temperature (°C)	Thermal rating (amperes)
Ibis	397.5	100	640
Drake	795	100	995
Falcon	1590	100	1520
Drake/ACSS	795	200	1600

Calculations leading to optimization plots such as shown in Fig. 14-1 are usually done assuming a relatively simple line model consisting of a conductor in catenary between structures at a typical spacing.[4] In this "typical span" model, the line is approximated as a series of structures that have the same height and spacing so that the conductor between them has the same sag and tension in all spans. Typical numbers of angle and dead-end structures are assumed per mile of line. Structure height is just sufficient to meet ground clearance, and structure cost is estimated based on this height, on phase spacing, and on typical transverse, vertical, and longitudinal loads for this span. Such a simple typical span model yields exact electrical losses, approximate structure costs, and is adequate for the exact calculation of radio noise, audible noise, and electric and magnetic fields.

Having used the "typical span" model to determine the range of conductor sizes which yield minimum total present worth cost of electrical losses and construction costs and adequately low environmental effects, the transmission-line design can be further optimized by considering a more realistic "terrain optimized" model of the line on actual or typical terrain. In such a study, the designer utilizes the availability of fast, efficient tower spotting algorithms to provide more exact structure cost estimates. Such studies have been described in Refs. 5 and 6.

Optimization of transmission designs using modern computer-based techniques allows the designer to consider variations in standard design constraints by modeling alternate designs having various design constraints on the same terrain. For example, transmission-line designs normally assume a standard unloaded conductor tension. Optimization studies might include evaluation of higher than standard conductor tensions in order to reduce conductor sag at high temperature. A "typical span" model may be used to evaluate the savings in structure height due to reduced sag and the increased cost of angle structures due to higher tension levels. A "terrain optimized" model will provide a more realistic estimate of the savings in structure height and the increased cost of angle structures and dead ends and will also identify costs related to uplift of structures at minimum temperature.

In addition to conductor tension, a "terrain optimized" model of the proposed line allows the designer to estimate costs for variations in

Available structure classes (e.g., fewer tangent types, an added light angle structure)

Conductor type (e.g., percentage of steel area in ACSR, self-damping conductor)

Available structure heights (e.g., fewer available heights, taller structures)

Optimization studies involve the consideration of nonstandard conductors and structures. This is typically justified only by large-scale design and construction projects or during the development of new standard transmission designs to meet changes in environmental effect constraints. Reuse of existing "standard" structure designs or conductors is often preferred due to considerations such as spare parts, tools and training, maintenance, known reliability, externally imposed factors such as hot line maintenance clearances, and short or highly constrained construction.

A highly variable component of transmission line costs is getting permits and rights-of-way. In some extreme situations this may be so great as to counter balance the normally much higher cost of underground cables.

14.1.4 Electrical Properties of Conductors

Positive-Sequence Resistance and Reactances. The conductors most commonly used for transmission lines have been aluminum conductor steel-reinforced (ACSR), all-aluminum conductor (AAC), all-aluminum alloy conductor (AAAC), and aluminum conductor alloy-reinforced (ACAR), but conductors able to operate at higher temperatures such as ACSS are available for a modest price premium and are becoming more common. Research is progressing on new high temperature ceramic-cored conductors. Tables of the electrical characteristics of the most commonly used ACSR conductors are in Sec. 4. Characteristics of other conductors can be found in conductor handbooks or manufacturers' literature and web sites.

The per mile resistance, inductive reactance, and capacitive reactance can be determined from the data in the tables of Sec. 4 and the spacing factors X_d and X_d.

The positive-sequence resistance is listed as the 60-Hz value at 50°C. The expression for inductive reactance per mile is

$$X_L = 0.004657f \log \frac{D}{\text{GMR}} \tag{14-4}$$

where D = equivalent spacing in feet, GMR = geometric mean radius in feet as given in the conductor tables of Sec. 4, and f = frequency in hertz. GMR for ACSR conductor is given at 60 Hz. However, 60-Hz values of GMR can be used at other commercial power-system frequencies with small error. X_L also can be expressed as

$$X_L = X_a + X_d = 0.004657f \log \frac{1}{\text{GMR}} + 0.004657f \log D \tag{14-5}$$

When the spacing is 1 ft, X_d becomes zero. Thus X_d is frequently called the "one-foot" inductive reactance. The expression for capacitive shunt reactance per mile is:

$$X_c = \frac{4.099 \times 10^6}{f} \log \frac{D}{r_c} \tag{14-6}$$

where r_c = conductor radius in feet, which can also be expressed as

$$X_c = X'_a + X'_d$$

where

$$X'_a = \frac{4.099 \times 10^6}{f} \log \frac{1}{r_c} \tag{14-7}$$

and

$$X'_d = \frac{4.099 \times 10^6}{f} \log D \tag{14-8}$$

Bundle conductors consist of two or more conductors per phase mechanically and electrically connected and supported by an insulator assembly. The positive-sequence resistance is, to a first approximation, the 60-Hz, 50°C values in the Sec. 4 tables divided by the number of conductors per phase. General formulas for the inductance and capacitance of bundle conductors are

$$L_\phi = \frac{1}{n}\left[0.74113 \log \frac{r_c}{\text{GMR}} + 0.74113 \log \frac{24(S_{gm})^n}{d(M_{gm})^{n-1}}\right] \quad \text{mH/mi} \tag{14-9}$$

From Eq. (14-9) inductive reactance is found to be

$$X_L = \frac{1}{n}\left[K + 0.004657f \log \frac{24(S_{gm})^n}{d(M_{gm})^{n-1}}\right] \quad \Omega\text{/mi at 60 Hz} \tag{14-10}$$

and the capacitance is

$$C_\phi = \frac{0.03883n}{\log[24(S_{gm})^n/d(M_{gm})^{n-1}]} \quad \mu\text{F/mi} \tag{14-11}$$

In the above, n = number of conductors per phase (bundle); d = diameter of conductor in inches; S_{gm} = geometric mean distance between conductors of different phases in feet, found by taking the mean distance from all conductors of one phase to all conductors of the other phases; M_{gm} = geometric mean distance in feet between the n conductors of one phase; K = internal conductor reactance defined as

$$K = 0.004657 f \log \frac{r_c}{\text{GMR}} \qquad \Omega/\text{mi} \tag{14-12}$$

TABLE 14-2 Equivalent Reactances

Bundle	X_{aeq}	X'_{aeq}
2 conductors	$^1/_2(X_a - X_s)$	$^1/_2(X'_a - X'_s)$
3 conductors	$^1/_3(X_a - 2X_s)$	$^1/_3(X'_a - 2X'_s)$
4 conductors	$^1/_4(X_a - 3X_s)$	$^1/_4(X'_a - 3X'_s)$

The inductive series reactance and capacitive shunt reactances for bundled conductors can also be found by using the $X_a + X_d$ method, by determining the equivalent X_a and X'_a of the conductor bundle. The expressions for the equivalents are given in Table 14-2. These expressions are for three-conductor bundles on equilateral spacing and for four-conductor bundles on square spacing. The subscript s indicates the spacing of the conductors within the bundle in feet. Values for X_a and X'_a are in the conductor tables in Sec. 4. Values for X_s and X'_s are from the same formulas as X_d and X'_d.

$$X_s = 0.004657 f \log s \tag{14-13}$$

$$X'_s = \frac{4.099 \times 10^8}{f} \log s \tag{14-14}$$

where s is in feet and f is frequency in hertz. Equation (14-14) is correct for a ratio of spacing s to conductor radius r of 5 or more.

The value of X_{aeq} is added to X_d (the spacing factor, which is determined for the mean spacing between the conductors of the different phases). X'_{aeq} and X'_d are handled in a like manner.

Zero-Sequence Impedances. When earth-return currents due to faults or other causes are to be calculated, negative- and zero-sequence impedances must be determined in addition to positive-sequence quantities. Negative-sequence quantities are the same as the positive-sequence values for transmission lines. Precise determination of the zero-sequence quantities is difficult because of the variability of the earth-return path.

Calculation of zero-sequence impedance parameters is far more complex than for positive-sequence quantities, being a function of conductor size, spacing, relative position of conductors with respect to overhead ground wires, electrical characteristics of overhead ground wires, and the resistivity of the earth-return circuit. Reference 7 includes a detailed analysis of zero-sequence parameters, which are normally calculated using digital computer programs.

Table 14-3 lists representative values of positive- and zero-sequence impedances for different voltage transmission lines with shield wires. Zero-sequence reactance increases for unshielded lines.

TABLE 14-3 Typical Transmission-Line Impedance*

Voltage, kV	R_1	X_{L1}	X_{C1}	R_0	X_{L0}	X_{C0}	X_0/X_1
69	0.280	0.709	0.166	0.687	2.74	0.315	3.86
115	0.119	0.723	0.169	0.625	2.45	0.265	3.39
230	0.100	0.777	0.182	0.591	2.26	0.275	2.91
345	0.060	0.590	0.138	0.551	1.99	0.208	3.37
500	0.028	0.543	0.127	0.463	1.90	0.198	3.50
765	0.019	0.548	0.128	0.428	1.77	0.185	3.23

*R_1, X_{L1}, R_0, X_{L0} are in ohms per mile; X_{C0}, X_{C1} are in megohm-miles.
Note: 1 mi = 1.61 km.

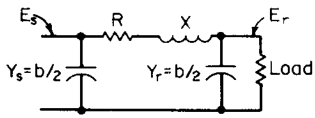


FIGURE 14-2 Nominal-π line.

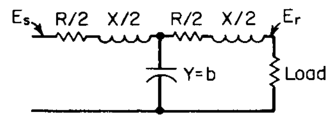


FIGURE 14-3 Nominal-*T* line.

Nominal-π Representation. Transmission lines can be represented by nominal π as in Fig. 14-2, in which half the capacitive susceptance, in siemens, is connected at each end of the line. The nominal-π representation is used in digital computer studies involving lines of moderate length (usually under 100 mi).

Nominal-T Representation. The nominal-T representation of a transmission line is shown in Fig. 14-3. The total line susceptance b, in siemens, is concentrated at A, the midpoint of the line.

ABCD Parameters. These line parameters (general circuit constants) are defined by the equations

$$E_s = AE_r - BI_r \tag{14-15}$$

$$I_s = CE_r - DI_r \tag{14-16}$$

For a short line (under 100 mi) if $Z_1 = R + j\omega L$ and $Z_2 = 2/jb$ (refer to the nominal-π line of Fig. 14-2)

$$A = D = \frac{Z_1 + Z_2}{Z_2} \tag{14-17}$$

$$B = Z_1 l \tag{14-18}$$

$$C = \left(\frac{Z_1 + 2Z_2}{Z_2^2}\right)/l \tag{14-19}$$

For longer lines where l is the length of the line

$$A = D = \cosh(\gamma l) \tag{14-20}$$

$$B = Z_c \sinh(\gamma l) \tag{14-21}$$

$$C = \frac{\sinh(\gamma l)}{Z_c} \tag{14-22}$$

where

$$\gamma = \sqrt{(R + j\omega L)(j\omega C)} \tag{14-23}$$

and

$$Z_c = \sqrt{\frac{R + j\omega L}{j\omega C}} \tag{14-24}$$

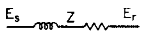
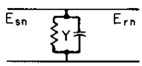
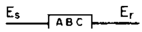
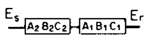
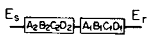
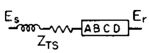
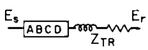
and R, L, and C are line resistance, inductance, and capacitance per mile.

Formulas for *ABCD* constants for various circuit configurations are given in Table 14-4.

Surge Impedance Loading. The surge impedance of a transmission line is the characteristic impedance with resistance set equal to zero (i.e., R is assumed small compared to $j\omega L$ of Eq. 14-24).

$$Z_s = \sqrt{\frac{L}{C}} \tag{14-25}$$

TABLE 14-4 Formulas for Generalized Circuit Constants

			Equivalent constants			
No.	Type of network		A_t	B_t	C_E	D_t
1	Series impedance		1	Z	O	1
2	Shunt admittance		1	O	Y	1
3	Uniform line		A	B	C	A
4	Two uniform lines		$A_1A_2 + C_1B_2$	$B_1A_2 + A_1B_2$	$A_1C_2 + A_2C_1$	$A_1A_2 + B_1C_2$
5	Two nonuniform lines or networks		$A_1A_2 + C_1B_2$	$B_1A_2 + D_1B_2$	$A_1C_2 + D_2C_1$	$D_1D_2 + B_1C_2$
6	General network and sending transformer impedance		$A + CZ_{TS}$	$B + DZ_{TS}$	C	D
7	General network and receiving transformer impedance		A	$B + AZ_{TR}$	C	$D + CZ_{TR}$
8	Two networks in parallel		$\frac{A_1B_2 + A_2B_1}{B_1 + B_2}$	$\frac{B_1B_2}{B_1 + B_2}$	$C_1 + C_2 + \frac{(A_1 - A_2)(D_2 - D_1)}{B_1 + B_2}$	$\frac{D_1B_2 + D_2B_1}{B_1 + B_2}$

Note: All constants in this table are complex quantities; $A = a_1 + ja_2$ and $D = d_1 + jd_2$ are numerical values, $B = b_1 + jb_2$ = ohms, and $C = c_1 + jc_2$ = siemens. As a check on calculations of *ABCD* constants, note that $AD - BC = 1$.

The power which flows in a lossless transmission line terminated in a resistive load equal to the line's surge impedance is denoted as the surge impedance loading (SIL) of the line. Under these conditions, the receiving end voltage E_R equals the sending end voltage E_S in the magnitude, but lags E_S by an angle δ corresponding to the travel time of the line. For a 3-phase line

$$\text{SIL} = \frac{(E_{L\text{–}L})^2}{Z_S} \tag{14-26}$$

Since Z_s has no reactive component, there is no reactive power in the line, $Q_S = Q_R = 0$. This indicates that for SIL the reactive losses in the line inductance are exactly offset by reactive power supplied by the shunt capacitance or $I^2\omega L = E^2\omega C$.

SIL is a useful measure of transmission-line capability even for practical lines with resistance, as it indicates a loading where the line's reactive requirements are small. For power transfer significantly above SIL, shunt capacitors may be needed to minimize voltage drop along the line, while for transfer significantly below SIL, shunt reactors may be needed.

SILs for typical transmission lines are given in Table 14-5. Cables normally have current ratings (ampacity) considerably below SIL, while overhead line current ratings may be either greater than or less than SIL. Figure 14-4 presents illustrative overhead line loadability as a function of line length and SIL.

Although Fig. 14-4 is illustrative only of loading limits, it is a useful estimating tool. Long lines tend to be stability-limited and have a lower loading limit than shorter lines, which tend to be voltage-drop- or conductor-ampacity-limited.

TABLE 14-5 SIL of Typical Transmission Lines

System kV	Z_s, Ω	SIL, MW
	Overhead lines	
230	367	144
345	300	400
500	285	880
65	280	2090
1200	250	5760
	Cables	
230	38	1390
345	25	4760

14.1.5 Electrical Environmental Effects

Corona and Field Effects. There are two categories of electrical environmental effects of power transmission lines. Corona effects are those caused by electrical stresses at the conductor surface which result in air ionization ("corona") and include radio, television, and audible noise. Field effects are those caused by induction to objects in proximity to the line. While the generic term is electromagnetic effects, within the electric power industry the fields are divided into two types: electric-field effects and magnetic-field effects. Electric fields, related to the voltage of the line, are the primary cause of induction to vehicles, buildings, and objects of comparable size. Magnetic fields, related to the currents in the line, are the primary cause of induction to long objects, such as fences and pipelines.

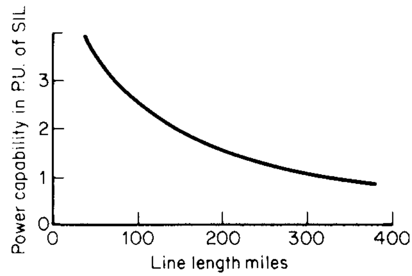


FIGURE 14-4 Overhead line loading in terms of SIL.

Assessment Criteria. In an electrical environmental analysis, it is important to determine the proper criteria for assessment of the impact. For example, the audible noise criterion in a commercial or industrial area would be inappropriate in a quiet residential neighborhood.[8] Likewise, ground-level electric field criteria on a parking lot would be different from that in terrain inaccessible by motor vehicles. For audible noise, the only concern is annoyance, but for electric fields, safety, annoyance, and perception levels all may have to be considered.

Probability of exposure is also an important criterion. The impact of radio noise in arid locations is different from that in places with considerable rainfall. Since different people have different perception and annoyance thresholds, statistical evaluations are necessary, recognizing that some percentage of people will find a generally accepted noise level annoying. Because of the combination of worst-case events which are normally assumed in an electrical environmental analysis, the overall probability of annoyance is usually considerably smaller than initially presumed.

A predictive model is necessary to calculate the expected effect. Depending on the specific effect, it may be an empirical formula or may be quite sophisticated. However, it is only by calculating the effect and comparing it with specified criteria that the overall impact can be assessed. This is illustrated by Fig. 14-5,[9] which is a flowchart of the analysis procedure for an example case of electric-field-induced shock.

Audible Noise. Corona-produced audible noise during foul weather, particularly during or following rain, can be an important design parameter for high-voltage ac transmission lines. Audible noise has two components, a random noise component and a low-frequency hum, each produced by different physical mechanisms. While the hum component is closely correlated with corona loss on the line, the random noise is not. Of these two, the most frequent cause of annoyance is the random noise, and it is this which is calculated and compared with acceptance criteria.

Analyses to predict levels of audible noise consider A-weighted sound level [dB(A)] during rain, including

L_{50}, which is the level exceeded 50% of the time during rain (considering all rain storms over a period of time, usually 1 year).

L_5, which is the level exceeded 5% of the time during rain.

Average, which is the average level of noise expected during rain. (This is usually close to the L_{50} value and is sometimes called "wet-conductor" noise.)

Heavy rain, which is the level expected during heavy rain. (This usually is representative of laboratory artificial rain tests but is assumed representative of the L_5 level.)

Reference 10 compares audible noise formulas, which have been developed throughout the world. One formula for both L_5 and L_{50} values is given by

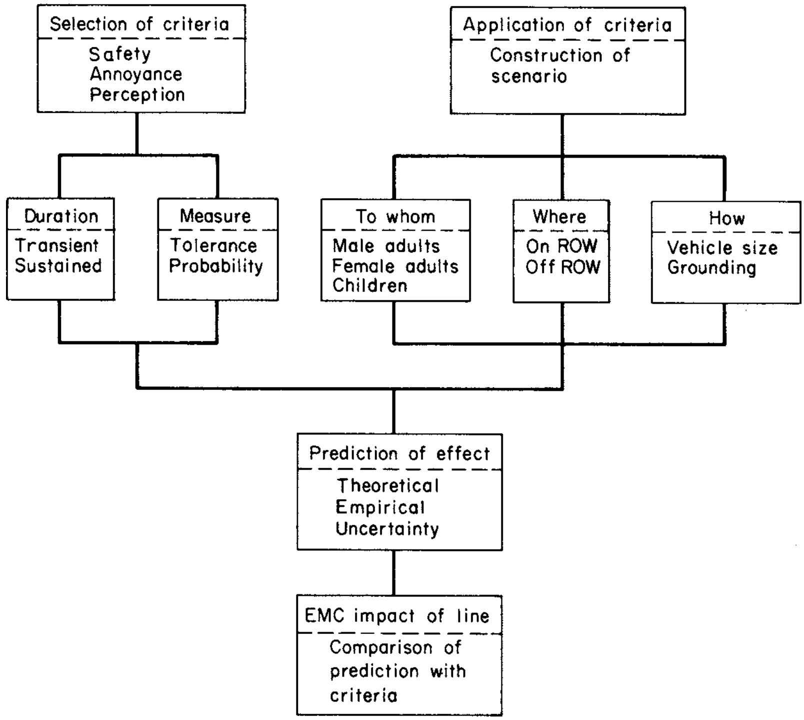


FIGURE 14-5 Factors affecting transmission line EMC for shock effects.

g = Average-maximum surface gradient of conductor or conductor bundle, kV/cm

n = Number of subconductors in a phase (or pole) bundle

d = Diameter of subconductors, cm

D = Distance from line to point at which noise level is to be calculated, m

SL = A-weighted sound level of the noise produced by the line, dB(*A*)

N_p = number of phases

AN = A-weighted sound level of the noise produced by one phase of the line, dB(*A*)

AN_0 = A reference *A*-weighted sound level, dB(*A*)

K_1, K_2, K_3, K_4 = Constant coefficients

Application = All line geometries

Noise measure = L_5 rain and L_{50} rain

Range of validity = 230–1500 kV, $1 \leqslant n \leqslant 16$, $2 \leqslant d \leqslant 6$

For each phase, the L_5 noise level is given by

$$AN_5 = \frac{-665}{g} + 20 \log n + 44 \log d - 10 \log D - 0.02\, D + \text{AN}_0 + K_1 + K_2 \qquad \text{(14-27)}$$

with

$$\begin{aligned} \text{AN}_0 &= 75.2 && \text{for } n < 3 \\ &= 67.9 && \text{for } n \geq 3 \\ K_1 &= 7.5 && \text{for } n = 1 \\ &= 2.6 && \text{for } n = 2 \end{aligned}$$

$$K_2 \begin{aligned} &= 0 && \text{for } n \geq 3 \\ &= 0 && \text{for } n < 3 \\ &= 22.9(n-1)\frac{d}{B} && \text{for } n \geq 3 \end{aligned}$$

where B is the bundle diameter, cm.

The L_{50} level for each phase is obtained from

$$\mathrm{AN}_{50} = \mathrm{AN}_5 - \Delta A \tag{14-28}$$

where

$$\begin{aligned} \Delta A &= 14.2\frac{g_c}{g} - 8.2 && \text{for } n < 3 \\ &= 14.2\frac{g_c}{g} - 10.4 - 8\left[(n-1)\frac{d}{B}\right] && \text{for } n > 3 \end{aligned}$$

and

$$\begin{aligned} g_c &= 24.4\,(d^{-0.24}) && \text{for } n \leq 8 \\ &= 24.4\,(d^{-0.24}) - 0.25\,(n-8) && \text{for } n > 8 \end{aligned}$$

$$\mathrm{SL} = 10 \log \sum_{i=1}^{Np} 10^{AN_i/10} \tag{14-29}$$

Figure 14-6 illustrates a typical presentation of audible noise calculations. The profile, in this case for a representative 500-kV line and wet conductors, quantifies the level of noise in dB(A) greater than 0.002 μbar as a function of distance from the centerline of the structure. From this method of presentation, analysis of maximum levels as well as effect on width of right-of-way can be analyzed. Similarly, design variables such as conductor size, spacing, and configuration; height of conductors; and weather variations can be considered.

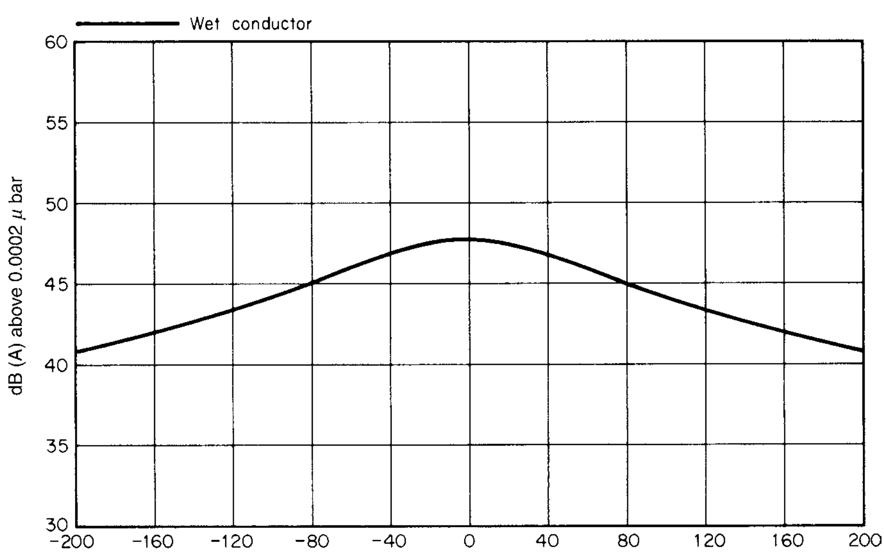


FIGURE 14-6 Audible noise profile at ground level for a transmission line.

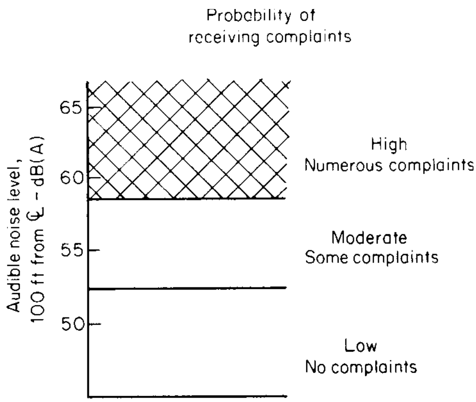


FIGURE 14-7 Audible noise compliance guidelines.

Figure 14-7[3,11] quantifies experience with transmission-line audible noise complaints. These occur mostly during wet-conductor conditions and low ambient noise, such as after rain or during fog. During heavy-rain conditions, the noise of the rain masks the line noise. Other factors during heavy rain, such as closed windows, combine to make this condition less likely to result in complaints even though the noise is louder. In the absence of local noise regulations, comparison of calculated L_{50} or average audible noise with Fig. 14-7 gives a reasonable preliminary evaluation of the possibility of audible noise annoyance. When measurements are to be taken to confirm ambient noise or line noise, care must be taken to follow proper procedures.[12]

Radio and Television Noise. Electromagnetic interference from overhead power lines is caused by two phenomena: complete electrical discharges across small gaps (microsparks) and partial electrical discharges (corona). Gap-type sources occur at insulators, line hardware, and defective equipment and are a construction and maintenance problem rather than a design consideration. They are responsible for about 90% of radio noise complaints and can be located and eliminated as they occur.[13] Conductor and hardware corona is considered during the design phase. On a properly designed line, conductor corona noise rarely results in television interference complaints except perhaps in weak signal fringe areas.

The specification of "corona-free" hardware is important to eliminate electromagnetic interference from conductor support hardware, and is especially important as lines are constructed with closer spacings and resulting higher electric fields on the hardware. Conductor clamps and other fittings, which were formerly acceptable at traditional phase spacings, may not be adequate for compact lines.

For ac lines, radio and television noise are functions of the weather. Fair-weather noise may be significant and varies with the season, wind velocity, and barometric pressure.

Two families of computation methods are available for radio noise: those based on conductor laboratory tests and analytical propagation theory (semianalytical methods) and those based on an empirical formula using data from long-term tests on operating lines (comparative methods).

The comparison method[14] is useful for conventional geometries and designs:

$$\mathrm{RI} = -150.4 + 120 \log g + 40 \log d + 20 \log \frac{h}{D^2} + 10\,[1 - (\log 10 f)^2] \qquad (14\text{-}30)$$

where g = average maximum surface gradient of conductor or conductor-bundle, kV peak/cm
d = subconductor diameter, mm
h = height of phase, m
D = radial distance to observer, m
f = frequency, MHz
RI = fair-weather radio noise, dB

RI is calculated for each phase and the maximum value is used as the RI of the line. Average foul-weather RI levels are assumed to be 17 dB above fair weather, and heavy-rain RI 24 dB above fair weather. Other methods are described in Ref. 3.

As with audible noise, the most useful data presentation is the level of radio noise as a function of distance from the centerline of the structure. An illustrative example for a specific 500-kV line is shown in Fig. 14-8.

There are no generally accepted RI limits in the United States, because of the impossibility of setting universal criteria for all land use and local conditions.[15] A Canadian standard exists for RI limits and is a useful guide.[16]

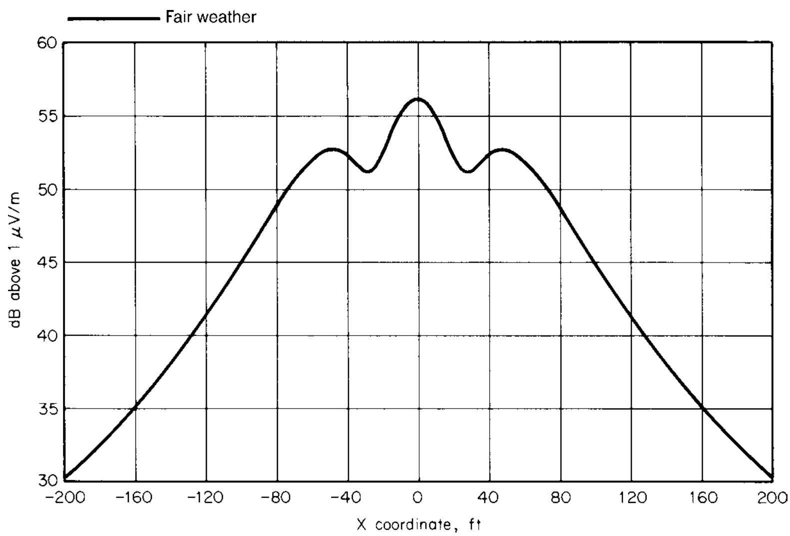


FIGURE 14-8 Radio noise profile at ground level for transmission line.

Two quantities are required to set criteria for evaluation of radio noise. These are the level of signal strength in the line vicinity and an appropriate signal/noise ratio. This latter ratio is typically assumed to be 24 to 26 dB at the edge of the right-of-way. Primary signal strengths may be 54 dB above 1 μV (0.5 mV/m) in rural areas to 88 dB or more in cities.

Prediction of television noise is not as advanced as that of radio noise, primarily because of the limited number of actual cases of conductor corona television interference. As with radio noise, most television interference complaints result from microsparks which can be located and eliminated as they occur. These are not generally a design consideration. In the few cases where corona-caused television noise has occurred in foul weather, it has often been possible to remedy the situation by an improvement in the receiving antennas rather than changes to the transmission-line design. References 3 and 17 contain recent work on prediction and evaluation of TVI.

***Gaseous Oxidants*.** Gaseous oxidants can be produced by corona activity in air and, in sufficient concentrations, may produce adverse effects on flora and fauna. The most important oxidants are ozone (O_3) and oxides of nitrogen (mainly NO and NO_2), where ozone is the major constituent.

Federal standards limit photochemical oxidants to 0.12 part per million for a maximum of 1-h concentration not to be exceeded more than once per year. Some states have more restrictive regulation; for example, the Minnesota Pollution Control Agency standards are for 0.07 ppm by volume (130 μg/m^3). Ozone can be detected by smell at minimum concentrations of 0.01 to 0.15 ppm.

Analytic studies and field measurements have been conducted on both operating and test lines.[18–25] The highest calculated value for 1-mi/h wind parallel to the line was 0.019 ppm maximum ground-level concentration. Measurements have indicated that transmission-line contribution to gaseous oxidants cannot be detected within statistical limits of significance and accuracy. With instrumentation capable of detecting 0.002 ppm, the transmission-line contribution was indistinguishable from ambient.

Thus, gaseous oxidants are not a concern with respect to electric power transmission lines.

Ground-Level Electric Fields. Ground-level electric field effects of overhead power transmission lines relate to the possibility of exposure to electric discharges from objects in the field of the line. These may be steady currents or spark discharges. Other areas which have received attention are the possibility of fuel ignition and interference with wearers of prosthetic devices (e.g., pacemakers).[26]

It is appropriate to consider unlikely conditions when setting and applying electric-field safety criteria because of possible consequences; thus statistical considerations are necessary. Annoyance

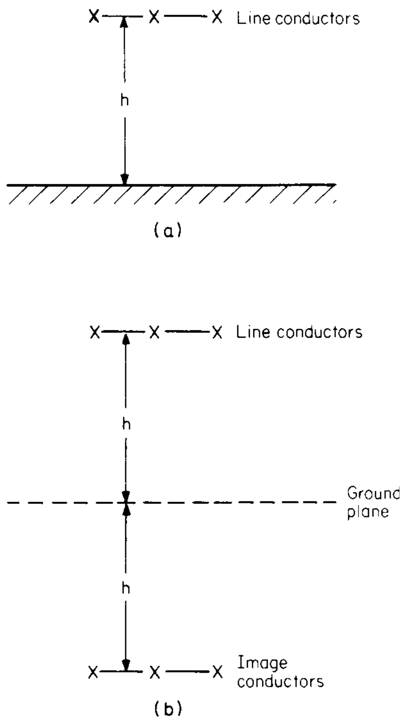


FIGURE 14-9 Representation of conducting earth: (*a*) earth; (*b*) image.

criteria need not be as stringent and mitigating factors can be considered.

Electric-Field Calculations. The resultant electric fields in proximity to a transmission line are the superposition of the fields due to the three-phase conductors. The conducting earth must be represented by image charges located below the conductors at a depth equal to the conductor height.

For example, consider the three-conductor line of Fig. 14-9. The effect of earth can be represented by replacing the earth with image conductors as shown in Fig. 14-9. At 60 Hz and for typical values of earth resistivity, the relaxation time of the earth (the time required for charges to redistribute themselves due to an externally applied field) is so small compared to the power frequency wave that for each instant of time the charge is distributed on the earth's surface as in the static condition (i.e., the earth appears to be a perfect conductor).

The electric fields surrounding the transmission line are a function of the instantaneous charges on the line. Usually, however, the charges are not known, but the voltages to ground of the different conductors are. Since the charge Q on each conductor is a function of the voltage on all conductors, an $n \times n$ capacitance matrix results, where n is the number of conductors, according to the formula

$$[Q] = [C][V] \tag{14-31}$$

which, for a three-conductor configuration (ignoring shield wires), is

$$Q_1 = C_{11} V_1 + C_{12}V_2 + C_{13}V_3 \tag{14-32}$$

$$Q_2 = C_{21} V_1 + C_{22}V_2 + C_{23}V_3 \tag{14-33}$$

$$Q_3 = C_{31} V_1 + C_{32}V_2 + C_{33}V_3 \tag{14-34}$$

The off-diagonal (mutual) capacitance terms significantly affect the final results. The individual terms of the capacitance matrix are computed by

$$C_{nm} = \left.\frac{Q_n}{V_m}\right| \quad \text{all other voltages} = 0 \tag{14-35}$$

where n and m are conductors.

The potential coefficient matrix is, however, more amenable to computation and is defined by

$$[V] = [P][Q] \tag{14-36}$$

whose individual terms are given by

$$P_{nm} = \left.\frac{V_n}{Q_m}\right| \quad \text{all other charges} = 0 \tag{14-37}$$

This is an open-circuit matrix where the individual terms can be computed by assuming a charge at one conductor and calculating the voltage at the prescribed location assuming all the other conductors nonexistent (open-circuited). For a single conductor of radius r and a height h above the earth, the self-potential coefficient is given by

$$P_{nm} = \frac{1}{2\pi\epsilon_o} \ln \frac{2h}{r} \tag{14-38}$$

For two conductors n and m where d_{nm} is the distance between them, and $d_{nm'}$ is the distance between conductor n and the image of conductor m, the mutual potential coefficient is given by

$$P_{nm} = \frac{1}{2\pi\epsilon_o} \ln \frac{d_{nm'}}{d_{nm}} \tag{14-39}$$

This potential coefficient matrix can be calculated and inverted to yield the capacitance matrix:

$$[C] = [P]^{-1} \tag{14-40}$$

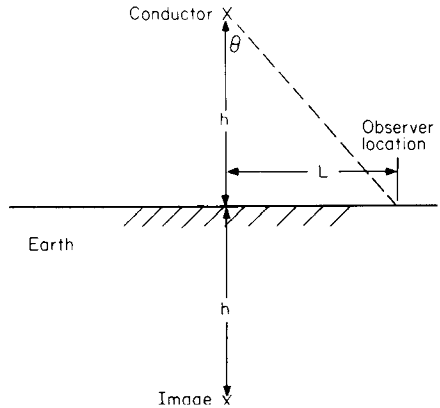


FIGURE 14-10 Single conductor.

This capacitance matrix allows the calculation of the charges on the individual conductors for the given initial voltage distribution according to Eqs. (14-32) through (14-34). Once these charges are obtained, the desired electric fields can be determined.

For the single conductor and observer location of Fig. 14-10, the ground-level electric field is determined from

$$E = \frac{Q_l}{2\pi\epsilon_o r} \tag{14-41}$$

The distance from the conductor to the observer is

$$r = \sqrt{h^2 + L^2} \tag{14-42}$$

Thus

$$E = \frac{Q_l}{2\pi\epsilon_o\sqrt{h^2 + L^2}} \tag{14-43}$$

Q must be determined from $[Q] = [C]\,[V]$. For a single conductor this equation reduces to

$$Q_l = P^{-1}V = \frac{1}{\frac{1}{(2\pi\epsilon_o)}\ln(2h/r)}V \tag{14-44}$$

For a multiconductor configuration, Q would come from the full matrix calculation.

E is radially directed from the line charge. The vertical component is

$$|E|\cos\theta = \frac{Q_l}{2\pi\epsilon_o\sqrt{h^2 + L^2}}\frac{h}{\sqrt{h^2 + L^2}} = \frac{Q_l}{2\pi\epsilon_o}\frac{h}{h^2 + L^2} \tag{14-45}$$

The vertical component of the electric field at ground level because of the image is equal to the field from the conductor, since the image is the geometric mirror image and has the opposite sign charge. Thus, the total ground-level field is given by

$$E = \frac{Q_l}{\pi\epsilon_o}\frac{h}{h^2 + L^2} \tag{14-46}$$

At ground level, the horizontal components of the electric fields of the conductor and its image cancel and the resultant field is purely vertical.

For a 3-phase line, the fields of the three conductors and their images are computed separately and added.

For fields extremely close to the line conductors, care must be taken to represent the local effects properly. For example, the surface field around the conductor is not uniform. For a bundled conductor, it is more nearly represented by a sinusoid. Farther from the conductors, a GMR representation will suffice.

For a bundle of diameter D with n conductors of radius r, the GMR is given by

$$\text{GMR} = \frac{D}{2}\sqrt[n]{\frac{2nr}{D}} \tag{14-47}$$

Replacing the conductor radius with the bundle GMR gives the appropriate representation.

Figure 14-11 illustrates a representative electric-field profile, in kV rms per meter, from the centerline of the structure. This presentation clearly illustrates the maximum field, the location of the maximum, and the effect on right-of-way width considerations. Sensitivity to various parameters can also be quickly evaluated.

Criteria for Evaluation. The effects of electromagnetic fields on humans is due to discharges from objects insulated from ground; typically vehicles, buildings, and fences which become electrically charged by induction from the line. Table 14-6 summarizes effects on humans, ranging from no perception through severe shock and possible ventricular fibrillation.[27]

Criteria for spark discharges are expressed in terms of stored charge or stored energy on the charged object. Levels for perception in adult males are of the order of 0.12 mJ, while experience indicates that approximately 2 mJ results in an annoying spark. Safety is seldom of concern, since approximately 25 J is required for injury, a value beyond that expected on objects beneath transmission lines.

Deno's work, using test data, relates short-circuit current to the undisturbed electric field for objects insulated from ground.[26] Initial calculations assume the worst possible combination of circumstances; no leakage path to ground exists for the object, complete grounding of the person involved, steady contact, and orientation of the vehicle parallel to the line. Table 14-7 lists sample criteria and electric fields needed to meet them for three sample vehicles.

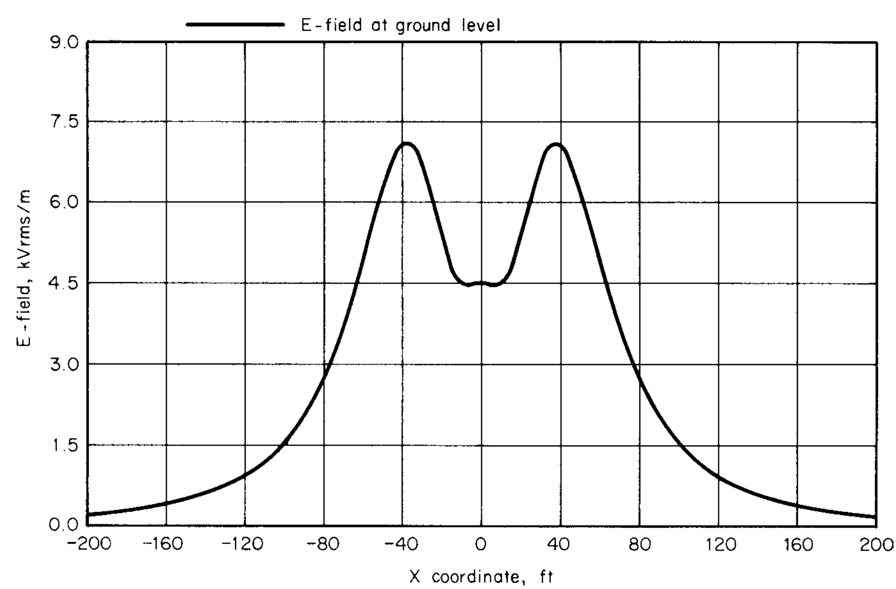


FIGURE 14-11 Electric-field profile at ground level for transmission line.

TABLE 14-6 Threshold Levels for 60-Hz Contact Currents

rms current, mA	Threshold reaction and/or sensation
	Perception
0.09	Touch perception for 1% of women
0.13	Touch perception for 1% of men
0.24	Touch perception for 50% of women
0.33	Grip perception for 1% of women
0.36	Touch perception for 50% of men
0.49	Grip perception for 1% of men
0.73	Grip perception for 50% of women
1.10	Grip perception for 50% of men
	Startle
2.2	Estimated borderline hazardous reaction, 50% probability for women (arm contact)
3.2	Estimated borderline hazardous reaction, 50% probability for women (pinched contacts)
	Let-go
4.5	Estimated let-go for 0.5% of children
6.0	Let-go for 0.5% of women
9.0	Let-go for 0.5% of men
10.5	Let-go for 50% of women
16.0	Let-go for 50% of men
	Respiratory tetanus
15	Breathing difficult for 50% of women
23	Breathing difficult for 50% of men
	Fibrillation
35	Estimated 3-s fibrillating current for 0.5% of 20-kg (44-lb) children
100	Estimated 3-s fibrillating current for 0.5% of 70-kg (150-lb) adults
	Established standards
0.50	ANSI standard for maximum leakage (portable appliance)
0.75	ANSI standard for maximum leakage (installed appliance)
5.0	NESC recommended limit for induced current under transmission line

TABLE 14-7 Limiting Electric Field for Given Criteria, kV/m

		Sample vehicles		
		Autos, pickups	Farm vehicles	Buses, trailer trucks
	Sample criteria	A	B	C
Safety	5 mA	22.32	10.86	6.33
	25 J	259.00	159.00	106.50
Annoyance	2 mA	8.92	4.35	2.50
	2 mJ	2.37	1.41	0.95
Perception	1.1 mA	4.91	2.39	1.39
	0.12 mJ	0.58	0.35	0.23

TABLE 14-8 Likely Range of Maximum Vertical Electric Field for Various Voltage Transmission Lines

Line voltage, kV	Near-ground vertical electric field, kV/m
765	8–13
500	5–9
345	4–6
230	2–3.5
161	2–3
138	2–3
115	1–2
69	1–1.5

High voltages may develop due to electric-field coupling, but the available short-circuit current is small (i.e., high-impedance source); thus calculations are based on a Norton equivalent and the short-circuit current. A relatively high resistance ground is sufficient to reduce electric-field-coupled voltage.

Table 14-8 lists maximum electric fields on the right-of-way under lines of different voltage classes. The fields attenuate rapidly with distance from the line and are usually much lower at the right-of-way edge.

Fuel Ignition. Theoretical calculations indicate that if several unlikely conditions exist simultaneously, a spark could release sufficient energy to ignite gasoline vapors. These conditions include a perfectly grounded person refueling a car perfectly insulated from ground with a metal can while the car is parked directly under a line. The spark would have to occur in the precise location of optimum fuel-air mixture. Research[3,28] confirms the low probability of accidental fuel ignition under actual conditions.

No confirmed cases of accidental ignition under transmission lines exist, confirming the low probability of these factors occurring simultaneously. Because of the consequences of a gasoline fire, some electric utilities advise that gasoline-fueled vehicles not be refueled near a line of 500 kV or above. If refueling were necessary, the vehicle could be grounded or the can connected to the vehicle to prevent sparks.

Ground-Level Magnetic Fields. Magnetic-field coupling affects objects which parallel the line for a distance, such as fences and pipelines, and is generally negligible for vehicle- or building-sized objects. As opposed to electric-field coupling, magnetic-field coupling is a low-voltage, low-impedance source with relatively high short-circuit currents. Single grounds are ineffective in preventing magnetically coupled voltages and multiple low-resistance grounds are needed. The resistance of the person touching a fence or pipeline is the dominant current-limiting impedance in the equivalent electrical circuit.[29] Calculations are based on a "longitudinal electromotive force" approach and are described in Refs. 30 to 32.

A consideration in the calculation of magnetic fields, which is different from the electric-field calculation, concerns the images. A perfectly conducting earth can be assumed for the electric-field problem, even for realistic values of earth resistivity. The assumption of a transmission line in free space (no earth at all) gives a closer approximation to the ground-level magnetic fields than does the assumption of a perfectly conducting earth for measurements near the line. At distances beyond 100 m, the effect of earth becomes increasingly more significant. The effect of conducting earth is frequently treated by use of an image conductor located at a greater depth in the earth than the conductors are above the earth. Distances of several hundred meters are commonly used for this image depth, according to the relation $D = 660\sqrt{\rho/f}$ meters where ρ is the earth resistivity in ohm-meters and f is the frequency. Magnetic-field calculations are given in Ref. 12, including the use of Carson's terms to evaluate the effects of imperfectly conducting earth.

It is normally adequate to consider conductors in free space without images. For the conductor of Fig. 14-8 without its image

$$B = \frac{\mu_o I}{2\pi r} = \frac{\mu_o I}{2\pi\sqrt{h^2 + L^2}} \tag{14-48}$$

This is then separated into vertical and horizontal components by multiplying by $\sin\theta$ and $\cos\theta$. In general, both components must be retained. For a 3-phase line, all conductors must be computed. Horizontal and vertical components of B from the three conductors must then be combined individually as phasors, considering the angles of the different currents. The combined horizontal and vertical components in general have different angles, causing their resultant to trace an ellipse in time. Single-axis magnetic-field meters with the sensing coil oriented for a maximum reading give the magnitude of the major axis of the field ellipse. A three-axis meter of the type presently used for data logging responds to the square root of the sum of the squares of the three field components (the "resultant" field). The resultant field can be as much as 41% greater than the major axis of the field ellipse for circularly polarized fields of the type which result from symmetrical conductor configurations.[33]

In the same manner, image currents at some assumed depth can be computed and their fields included. The use of matrix calculations allows inclusion of ground wires and bundled conductors as is the case of electric fields.

With both electric and magnetic fields it is essential to follow proper measurement procedures[33] for comparison with calculations.

For electric fields it is important that the field not be perturbed by the presence of the operator or other nearby objects. For both electric and magnetic fields, it is necessary to accurately know the conductor positions, the conductor height, the distance to the observer, and the line operating conditions (voltage and current). Magnetic-field measurements frequently differ from calculations for a number of reasons beyond errors in distance and clearance measurement:

1. Line current is continually varying, so in general it is not as well known as line voltage. In addition to uncertainty concerning the current magnitude at the time of the field measurement, line current unbalance in both magnitude and phase angle can be important. Unbalance has an increasingly significant effect on the magnetic field, the farther one moves from the line. Spot measurements, especially in homes and near distribution lines, are of limited usefulness to characterize exposure. For this reason, it is often advisable to statistically characterize the magnetic field. A statistical description of the field over time can be developed from measurements or calculations which assume balanced currents. It is also sometimes useful to develop a statistical distribution for a specific current level and an assumed maximum unbalance.
2. Related to current unbalance is circulating current in the shield wires, return currents in the earth, and currents in nearby pipes. These currents may cause significant differences between calculation and measurement.
3. The difference between single- and three-axis instruments has been described above. Two operators with different instruments can determine different answers based on the principles of measurement.
4. In nonuniform fields, such as around appliances, the size of the sensing coil and presence or absence of ferromagnetic core material will affect the reading of instruments equally well calibrated in a uniform field. Calibration must be made in a calibrating coil sufficiently large that the field is uniform over the area of the sensing coil, yet not so large that other nearby currents do not affect the field.
5. Harmonic currents have different effects depending on the frequency response of the instrument. Some instruments have a response linearly increasing with frequency, some are flat with frequency, and others have bandpass filters of different waveshapes.

14.1.6 Line Insulation

Requirements. The electrical operating performance of a transmission line depends primarily on the insulation. An insulator not only must have sufficient mechanical strength to support the greatest loads of ice and wind that may be reasonably expected, with an ample margin, but must be so designed as to withstand severe mechanical abuse, lightning, and power arcs without mechanically failing. It must prevent a flashover for practically any power-frequency operating condition and many transient voltage conditions, under any conditions of humidity, temperature, rain, or snow, and with such accumulations of dirt, salt, and other contaminants that are not periodically washed off by rains.[34]

Insulator Materials. The majority of present insulators are made of glazed porcelain. Porcelain is a ceramic product obtained by the high-temperature vitrification of clay, finely ground feldspar, and

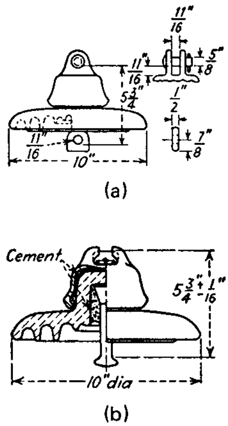


FIGURE 14-12 Typical porcelain disk insulator: (*a*) clevis type; (*b*) ball-and-socket type. (*Locke Insulators Inc.*)

silica. Insulators of high-grade electrical porcelain of the proper chemical composition free from laminations, holes, and cooling stresses have been available for many years.

The insulator glaze seals the porcelain surface and is usually dark brown, but other colors such as gray and blue are used. Porcelain insulators for transmission may be disks, posts, or long-rod types.

Porcelain insulators have been used at all transmission line voltages and, if correctly manufactured and applied, have high reliability.

A typical porcelain disk insulator is shown in Fig. 14-12.

Glass insulators have been used on a significant proportion of transmission lines. These are made from toughened glass, and are usually clear and colorless or light green. For transmission voltages they are available only as disk types. Most glass disk insulators will shatter when damaged, but without mechanically releasing the conductor. This provides a simple method of inspection.

Synthetic insulators, originally pioneered by the General Electric Company in 1963 for high-voltage transmission lines,[35] and more recently introduced by several manufacturers, are finding increasing acceptance. Most consist of a fiberglass rod covered by weather sheds of skirts of polymer (silicon rubber, polytetrafluoroethylene, cycloaliphatic resin, etc.)[36] as shown in Fig. 14-13. Other types include a cast polymer concrete called Polysil R[37] and a coreless type with alternating metal and insulating sections.[38]

Improvements in design and manufacture in recent years have made synthetic insulators increasingly attractive since their strength-to-weight ratio is significantly higher than that of porcelain and can result in reduced tower costs, especially on EHV and UHV transmission lines.

These insulators are usually manufactured as long-rod or post types. The light weight of most designs and resistance to damage aids construction. In addition, their performance under contaminated conditions may be significantly better than that of porcelain.[39]

Use of synthetic insulators on transmission lines is relatively recent and a few questions are still under study, in particular the lifetime behavior of insulating shed materials under contaminated conditions. It has been found necessary to use grading rings on some types at higher voltages to prevent damage to the sheds, and a very small number of insulators have experienced "brittle fractures," in which the fiberglass core breaks close to an end fitting. Despite these problems it appears that reliable synthetic insulators are presently available.

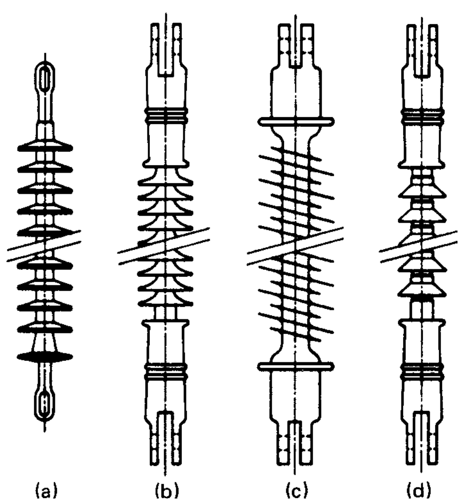


FIGURE 14-13 Typical nonceramic insulators.

Insulator Design. Transmission insulators may be strings of disks (either cap and pin or ball and socket), long-rods, or line posts. Posts are only infrequently applied above 230 kV.

Present suspension insulators conform to ANSI Standard C29.2, and standards have been established for 15,000-, 25,000-, 36,000-, and 50,000-lb ratings. It is common practice to use a factor of safety of 2 for the maximum mechanical stress applied to porcelain or glass insulators. For fiberglass-core insulators it is more common for the manufacturer to supply a recommended maximum working load.

Each manufacturer supplies catalogs which provide a physical description of the insulator's mechanical characteristics, wet and dry 60-Hz flashover strength, and positive- and negative-impulse (1.2×50 μs) critical (50%) flashover strength. Switching surge performance (250×3000 μs) is usually not supplied. In clean conditions most insulators of equivalent dimensions have very similar performance.

Suspension insulator strings, that is insulators used to support the conductor weight at a suspension or tangent structure, may be in I (vertical) or V configurations. The V configuration is used to prevent conductor movement and resultant clearance reductions at the structure. At dead-end or tension structures the insulators must also support the conductor tension, and it is not uncommon for these tension strings to be given a slightly higher flashover strength (e.g., by adding disks) to reduce the likelihood of a flashover that might lead to insulator string mechanical failure. Two or more strings of insulators in parallel can be used on suspension and tension strings to provide higher mechanical strength if required.

The electrical strength of line insulation may be determined by power frequency, switching surge, or lightning performance requirements. At different line voltages, different parameters tend to dominate. Table 14-9 shows typical line insulation levels and the controlling parameter. In compacted or uprated designs, considerably fewer insulators than these have been successfully used.[40,41]

Detailed descriptions of insulation design for electrical performance for different conditions, line voltages, and line types are available[42–44] from a number of studies.

Insulator Standards. The NEMA Publication *High Voltage Insulator Standards,* and AIEE Standard 41 have been combined in ANSI C29.1 through C29.9. Standard C29.1 covers all electrical and mechanical tests for all types of insulators. The standards for the various insulators covering flashover voltages; wet, dry, and impulse; radio influence; leakage distance; standard dimensions; and mechanical-strength characteristics are as follows: C29.2, suspension; C29.3, spool; C29.4, strain; C29.5, low- and medium-voltage pin; C29.6, high-voltage pin; C29.7, high-voltage line post; C29.8, apparatus pin; C29.9, apparatus post. These standards should be consulted when specifying or purchasing insulators.

Line Insulation Design. *Power-Frequency Design.* The criteria for power-frequency design is usually that flashover shall not occur for normal operating conditions, including reduced clearances to the structure from high wind. A typical wind-design limit is the 50- or 100-year return period wind, that is, a wind velocity which occurs only once in 50 or 100 years. This velocity is obtained from local wind records and may be typically 80 to 100 mi/h. Maximum operating voltages are designed by ANSI C84 and C92 standards and are 5% or 10% above the nominal value.

In clean conditions, power-frequency voltage is not a controlling parameter for insulator design (as distinct from air-gap clearance). However, even in quite lightly contaminated conditions it may become so.

Design for contamination is usually expressed as inches of creepage per kilovolt, where the creepage distance is the length of the shortest path for a current over the insulator surface and ranges up to 2 in/kV or more for heavy contamination. Standard insulator disks ($10 \times 5^3/_4$ in) have a typical creepage length of 11.5 in per disk. To avoid very long insulator strings for contamination, disks with additional creepage distance are made. The creepage can be extended by use of lengthened skirts and deeper grooves in the underside. Fog-type disks have up to 21.5 in of creepage per $13^1/_2 \times 8$-in units. A typical fog-type insulator is illustrated in Fig. 14-14.

TABLE 14-9 Typical Line Insulation

Line voltage, kV	No. of standard disks	Controlling parameter (typical)
115	7–9	Lightning or contamination
138	7–10	Lightning or contamination
230	11–12	Lightning or contamination
345	16–18	Lightning, switching surge, or contamination
500	24–26	Lightning, switching surge, or contamination
765	30–37	Switching surge or contamination

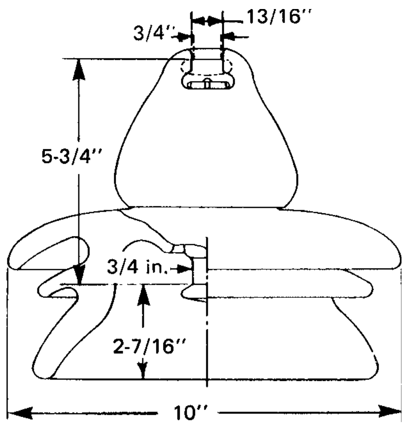


FIGURE 14-14 Typical fog-type disk insulator.

In extremely contaminated conditions, insulation with extended creepage may not be enough. In these cases insulator washing or the use of a silicone or petroleum grease coating (replaced at regular intervals) may be used.

Table 14-10 provides a simplified indication of creepage distance as a function of contamination,[42] and Fig. 14-15 shows guidelines from the IEEE application guide.[43]

For nonceramic insulation the same approach is used, except that subject to manufacturer's recommendations, a reduction in creepage distance up to 30% may be possible. This is due to the physical behavior of the nonceramic insulating material in moist conditions.

Another approach that has sometimes been used to combat contamination effects is the semiconductive glaze insulator. The semiconducting glaze allows a small but definite power-frequency current to flow over the surface. The insulator does not improve the standard test values, such as wet and dry power-frequency flashover and short-time impulse flashover, although it may have some value under switching surge conditions.

The glaze has a surface resistivity of about 10 MΩ per square. This is achieved by special formulations of materials involving, at the present stage of development, the use of tin-antimony additive to a more normal glaze composition. The presence of this small leakage current, of the order of 1 to 2 mA for suspension insulators, but which can be several times that value for large porcelains (such as are used in high-voltage bushings) has three effects:

1. Linearization of the voltage distribution over the insulator or string of insulators. This aids greatly in improving the performance of the insulator with respect to corona disturbance and RIV performance, plus having some benefits under dry and clean conditions.

TABLE 14-10 Insulation Requirements for Contamination: Provisional EHV Line Insulation Design Table for Various Contamination Conditions
Standard 5^5/34 × 10-in vertical insulator units

Contamination		Equivalent amount NaCl, mg/cm^2	Provisional design values		
			Leakage distance in/kV rms line to ground	Average kV rms	
Class	Types			Per in axial length	Per unit
A	Clean atmosphere—rural and forest regions; no industrial contamination	0–0.03	Insulation requirements not set by contamination		
B	Slight atmospheric contamination; suburbs of large industrial regions; railways; frequent washing rains	0.04	1.04	2.0	11.5
C	Moderate contamination containing soluble salts up to 5%; furnaces, dust from metallurgical plants, mine dust, fly ash, fertilizer dust in small quantities	0.06	1.31	1.6	9.1
D	Severe contamination containing 15% or more of soluble salts; dust from aluminum and chemical works, cement plants, heavy agricultural fertilizing, fly ash with high salt or sulfur content	0.12	1.74	1.2	6.9
E	Salt precipitation—seaside regions, salt marshes	0.30	2.11	1.0	5.7

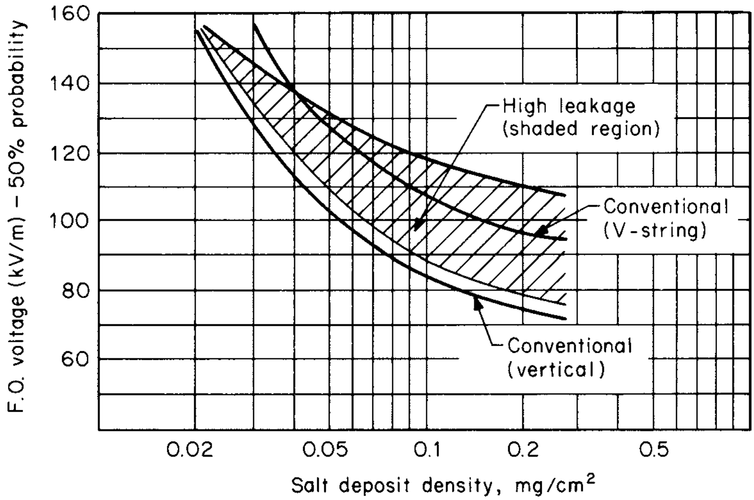


FIGURE 14-15 Power frequency withstand voltage of contaminated suspension insulators in fog expressed in kV/m of connection length (spacing).

2. Heating of the insulator. This occurs because of the power loss associated with the leakage current flow to a temperature which is usually about 5°C over the ambient air conditions. The heating effect enables the insulator to remain dry during conditions of fog or mist. This eliminates the majority of contaminated-insulator flashovers which occur when accumulated contamination becomes damp. This damp contamination condition is the most usual cause for contaminated-insulator flashover because most contaminants are more electrically conducting when damp or wet.

3. The elimination of “dry banding,” which is recognized as another major cause of flashover of standard insulators when contaminated. This occurs when the insulator has been thoroughly wetted, such as in a rain storm which wets but does not thoroughly clean the contamination from the insulator’s surface. Under these conditions dry bands will form as the standard insulator dries, and arcs strike across the dry-band area. These arcs can progress until flashover of the entire insulator occurs. With a semiconducting insulator, the relatively low resistance of the glaze shunts the dry-band area as the insulator dries and prevents the striking of the small power-frequency arcs.

The improved performance possible with semiconducting insulators has been proved in the laboratory and field,[45–49] but, because of the energy losses associated with the inherent leakage current, they are not widely used.

In some severe contamination areas, the problem has been effectively attacked by the use of silicone grease coatings. The unique amoebic action of a thick layer of silicone grease on an insulating surface is such as to envelop conducting solid particles which are said to “load” up the silicone grease to the saturation point, at which time the “used” silicone grease is removed and replaced with new silicone grease. In severe contamination areas, the greasing and degreasing cycles may be required every few months; in less severe contamination areas the cycle may be a year or more depending on experience acquired. In this manner, the time between insulator cleanings can be greatly extended, thus making for substantial savings. Once the silicone coating is used, the coatings must usually be wiped off and replaced manually, as necessary. Among the manufacturers of silicone grease are the General Electric Company and the Dow-Corning Corporation.

For the cleaning operation to remove contamination from the insulator surface, many contaminants such as salt deposits and water-soluble conducting liquids can be successfully removed by hot-line washing, using high-pressure water and insulated nozzles and hoses. Another method is "dry cleaning" by the use of an abrasive powder such as a limestone mixture or biodegradable plastic pellets, discharged at high pressure through hose and nozzle on the insulating surface. In many cases either hot-line washing or dry cleaning alone is sufficient to cope with the rate of accumulation encountered with the particular contaminant. An exception is substantially conducting materials, which take a chemical "set" after exposure to water, such as cement dust, some forms of gypsum, or asbestos, which often must first be manually chipped off or scrubbed off the insulating surface and then covered with silicone grease as previously described.

It should be emphasized that these problems may be very severe or even nonexistent, due to the variability of contamination exposure, which in turn depends on the chemical and electrical nature of the contaminant, prevailing wind direction, persistence of fog, smog, or other weather factors.

To monitor buildup of contaminants, some utilities collect data at the site to warn operating departments of an impending flashover, so as to promptly implement contamination-combative procedures.

Switching Surge Design. Operation of a circuit breaker on a transmission line can cause transient overvoltages, although flashovers due to such switching surges are rare in lines below 500 kV. If the breaker is opening, this may be due to restrikes across the breaker contacts as they separate, although restriking has been nearly eliminated with present breaker technology. If the breaker is closing, the cause may be unequal voltages on each side of the breaker, including the effect of residual charge on the line from a recent deenergization. The crest magnitudes of switching surges are normally defined in per unit of nominal power-frequency-crest phase-to-ground voltage. For example, on a 138-kV line (145 kV maximum), the per unit value is 118 kV. Typical switching surges range from 1 to as high as 4 or 5 per unit, and the varying characteristics of breaker operations provide a distribution of surge magnitudes which is often modeled as a truncated gaussian distribution.

The criterion for switching surge design is usually that flashover shall not occur for most or all switching events. Several design methods have been used, including

1. The maximum expected surge is determined, for example, from a transient network analyzer (TNA) or digital study, and the line insulation is designed to withstand that surge.
2. Rather than the maximum surge, a surge value corresponding to a statistical level is used, typically the 2% value (i.e., the crest value determined from the statistical distribution of surge crests, such that the level will be exceeded by only 2% of all surges).
3. Rather than design insulation to withstand a maximum surge, a statistical approach is used to design for a low number of flashovers per switching event. Typical levels are one flashover per 100 or 1000 breaker operations. This often results in a more economical design than either of the withstand approaches above.
4. By modeling the statistical distribution of switching surge crests, the distribution of insulator flashover with voltage, and the statistical distribution of weather that can be obtained from local weather stations, a probabilistic design can be prepared using a relatively simple computer program based on the allowable flashover rate. Typical procedures, data, and examples for such calculations are provided in several publications.[50,51]

Impulse Surge Design. Impulse surges on a line are caused by lightning strokes to or near the line. At transmission insulation levels, only strokes that directly intercept the line are capable of causing flashovers.

A number of methods of calculating transmission-line lightning performance have been published, and are summarized in the references to Chap. 12 of Ref. 3, together with a simplified calculation method. A computer program for this simplified calculation method is available from the *IEEE WG on Transmission Line Lightning Performance,* and more sophisticated programs for evaluation of multicircuit lines are available from a number of sources.

It is unusual for line insulation to be determined by lightning performance alone. More typically, insulation is determined by other requirements and the lightning performance is then verified. If this performance is unsatisfactory, it is often more efficient to change other design parameters such as shield wires or grounding than to add insulation.

Other methods of improving lightning performance have included addition of surge arresters at relatively frequent intervals along a line, and on double-circuit lines the use of unbalanced insulation so one circuit will flash over first and protect the other. Use of line arresters is most beneficial in regions of high ground resistance. Use of unbalanced insulation can improve the performance of the circuit with the highest insulation, but at the detriment of overall line performance.

Phase-to-Phase Insulation. The controlling paths for flashovers on most presently installed transmission lines are phase-to-ground, since there are usually grounded structure components between phases. However, for some new designs, such as the Chainette,[52] and compact lines the controlling path may be phase-to-phase air gaps or even phase-to-phase insulators.

Design methods for phase-to-phase insulation are essentially the same as for phase-to-ground insulation. Until recently, there was lack of knowledge of conductor clearance at midspan under various dynamic loading conditions, and lack of phase-to-phase switching surge data. Research studies sponsored by EPRI have now provided adequate design information on both topics.[44,50,52]

Protective and Grading Devices. Damage to insulators from heavy arcs was a serious maintenance problem in the past, and several devices were developed to ensure that an arc would stay clear of the insulator string. Subsequent improvements in the use of overhead ground wires and fast relaying have reduced the likelihood of insulator damage to the point that arc protection devices are now rarely used in the United States.

Earlier protective measures consisted of attaching small horns to the clamp, but it was found that horns with a large spread both at the top of the insulator and at the clamp were required to be effective. Under lightning impulse the arc tends to cascade the string, and tests show that the gap between horns should be considerably less than the length of the insulator string. Protection by arcing horns thus resulted in either a reduced flashover voltage or an increase in the number of units and length of the string. In any event, flashover persisted as a power arc until the line tripped out. For these reasons arcing horns have not been used in the United States for many years, although they are fairly common in Europe.

The arcing ring or grading shield is mainly for the purpose of improving the voltage distribution over the insulator string, and its effectiveness is due to the more uniform field. Protection of the insulator is not, therefore, dependent on simply providing a shorter arcing path, as is the case with horns. Efficient rings are rather large in diameter and, for suspension strings, clearances to the structure should be at least as great as from ring to ring. These considerations have made this device generally unattractive for modern construction. Grading rings are now used only at very high voltages for special applications, or with nonceramic insulators. Corona shields help improve the voltage distribution at the line ends of insulator strings.

14.1.7 Line and Structure Location

Preparation for Construction. The cost of preparing for transmission-line construction is a considerable part of the total costs—under some conditions as much as 25%. Right-of-way and clearing are more or less fixed by local conditions, but the cost of surveys, accompanying maps, profiles, and engineering layout is to some extent governed by judgment. Many times in the past the overall costs have been increased by right-of-way difficulties and by delays in receiving proper materials because of inadequate preparations. The engineering work, properly carried out, makes it possible to obtain the right-of-way and complete the clearing well in advance of construction and to purchase every item of material and deliver it to the correct location.

The work of locating and laying out a line does not require great refinement, but careful planning is essential. With inexperienced surveyors or drafters, it must be assumed that errors will be made, and every possible device must be used to discover these errors before construction is started.

Location. The general character of the line location should be determined because it has a definite bearing on the type of design. In extreme cases, such as difficult mountainous sections or in highly developed areas near cities, this may be a determining factor in the selection of the conductor and type of structures.

On heavy trunk lines, minor repairs and replacements are not an important item, and accessibility may often be rightly sacrificed to obtain the economy of a more direct route. Light wood lines must, however, be readily accessible for inspection and repairs. Line location is a matter of judgment and requires a person of wide general experience capable of correctly weighing the divergent requirements for inexpensive and available right-of-way, low construction costs, and convenience in maintenance. In mountainous country or in thickly populated areas, it is generally not advisable to attempt a direct route or try to locate on long tangents. Small angles of a few degrees cost little more and add little to the length of line. Most designs provide suspension structures for line angles of 5° to 15° which are not excessively costly. High, exposed ridges should be avoided, to afford protection against both wind and lightning.

Following a general reconnaissance by ground and air, for which 10 to 20 days per 100 mi should be allowed, and the assembling of all available maps and information, control points can be established for a general route or areas selected for more detailed study which may prove to be determining factors in the location of the line.

With this preliminary work completed, the major difficulties should have been determined. The policy as to such matters as right-of-way condemnation, electrical environmental assessments, telephone coordination, navigable-stream crossing, air routes, airports, and crossings with other utilities must be decided as definitely as possible.

Preliminary specifications should be issued before the final survey is started. These should include (1) outline drawings of the various structures with the important dimensions; (2) conductor sag curves and a sag template; (3) the maximum spans and angles for each type of structure; and (4) the requirements for right-of-way and clearing. Estimated costs are valuable, especially comparative costs of the various types of structure. With this information the field engineer can often, in a difficult section, choose the location best suited to the design.

Aerial maps can often be secured at much less cost than preliminary surveys, and in highly developed areas may be used to advantage for completely laying out the line without sending surveyors into the area until after the right-of-way has been secured.

Photographs taken at approximately 1/2 mi to the inch give sufficient detail for most work. Such maps can be photographically enlarged about four times for special detail. With a 1/2-mi-to-the-inch scale, the route of the line can be determined within a width of about 3 mi and sufficient landmarks located on a fairly accurate map to serve as a guide for flying the line.

Location Survey. The actual survey party can typically be divided into four divisions, each of which can complete at least a mile a day in average weather and country. Their operations may be carried out separately or nearly concurrently by allowing a full week's separation between successive operations and transferring personnel as needed.

The work falls naturally into the following: (1) an alignment party, choosing the exact location and cutting out the line; (2) a staking party, driving stakes at 100-ft stations and locating all obstructions; (3) a level party, taking elevations and side slopes; and (4) a property and topography party, locating property lines.

A field drafting force located at a convenient point for receiving field notes can complete the final plan and profile drawings as fast as the survey can be made.

The method of procedure and size of survey organization depend on the character of the country, the length and type of line, the experienced personnel available, and the schedule which must be maintained. In level, sparsely populated country, satisfactory but incomplete property surveys and profiles have been made during an open dry winter for a wood H-frame line 50 mi in length in approximately 4 months' time, with the personnel averaging a crew of eight and an engineer.

On a development involving the construction of several hundred miles of steel-tower line, the survey for a 65-mi line in rather difficult country, including 25 mi of inaccessible mountainous country, was completed with property maps and profiles in the form for permanent records in 2 months' time with a crew of about 20 and a locating engineer.

Purchase. Generally, right-of-way is not purchased in fee, but a perpetual easement is secured in which the owner grants the necessary rights to construct and operate the line but retains ownership and use of the land. The width of the right-of-way may be stated as a definite width or in general terms, but the easement must provide for (1) a means of access to each structure; (2) permission to erect all structures and guys; (3) all trees and brush to be cleared over a specified width for erection; (4) the removal of trees, which would not safely clear the conductor if the conductor were to swing out under maximum wind or which would not safely clear the conductor if they were to fall; and (5) the removal of buildings, lumber piles, haystacks, etc., which constitute a fire hazard. One of the major causes of serious line outages is the neglect to adhere strictly to conservative rules for clearing.

Tower Spotting. The efficient location of structures on the profile is an important component of line design. Structures of appropriate height and strength must be located to provide adequate conductor ground clearance and minimum cost. In the past, most tower spotting has been done manually, using templates, but several computer programs have been available for a number of years for the same purpose.

Manual Tower Spotting. A celluloid template, shaped to the form of the suspended conductor, is used to scale the distance from the conductor to the ground and to adjust structure locations and heights to (1) provide proper clearance to the ground; (2) equalize spans; and (3) grade the line (Fig. 14-16).

The template is cut as a parabola on the maximum sag (usually at 49°C) of the ruling span and should be extended by computing the sag as proportional to the square of the span for spans both shorter and longer than the ruling span. By extending the template to a span of several thousand feet, clearances may be scaled on steep hillsides. The form of the template is based on the fact that, at the time when the conductor is erected, the horizontal tensions must be equal in all spans of every length, both level and inclined, if the insulators hang plumb. This is still very nearly true at the maximum temperature. The template, therefore, must be cut to a catenary or, approximately, a parabola. The parabola is accurate to within about one-half of 1% for sags up to 5% of the span, which is well within the necessary refinement.

Since vertical ground clearances are being established, the 49°C no-wind curve is used in the template. Special conditions may call for clearance checks. For example, if it is known that a line will have high temperature rise because of load current, conductor clearance should be checked for the estimated maximum conductor temperature. One crossing over a navigable stream was designed for 88°C at high water. Ice and wet snow many times cause weights several times that of the 1/2-in radial ice loading, and conductors have been known to sag to within reach of the ground. Such occurrences are not normally considered in line design, and when they occur, the line is taken out of service until the ice or snow drops. Checks made afterward have nearly always shown no permanent deformation.

The template must be used subject to a "creep" correction for aluminum conductors. Creep is a nonelastic conductor stretch which continues for the life of the line, with the rate of elongation decreasing with time. For example, the creep elongation during the first 6 months is equal to that of the next 9 1/2 years. All conductors of all materials are subject to creep, but to date only aluminum conductors have had intensive study. Creep is not substantial in other conductors, but the conductor manufacturers should be consulted. The IEEE Committee Report, "Limitations on Stringing and Sagging Conductors," in the December 1964 *Transactions of the IEEE Power Group* discusses creep, and the reader should examine that report.[53]

Creep causes a continuous slow increase in the sag of the line which must be estimated and allowed for. The aluminum-conductor manufacturers will furnish creep-estimating curves, and most sag-tension computer programs now available are capable of calculating sags with and without creep. These curves are at approximately constant temperatures, around 15.5 to 21°C, and plot stress against elongation, one curve for each period of time, 1 h, 1 day, 1 month, 1 year, 10 years, etc. The values are integrated values for the period and are considered to be reasonable estimates. The temperature used is a reasonable average of the year's temperature across the center of the United States.

Precise values for creep are impossible to determine, since they vary with both temperature and tension, which are continuously varying during the life of the line. From Fig. 3 of the committee report in Ref. 53, it is found that a 1000-ft span of 954,000-cmil 48/7 ACSR when subjected to a constant tension of approximately 18% of its ultimate strength at a temperature of 15.5°C will have a sag increase in 1 day of approximately 5.5 in; in 10 days, 13 in; in 1 year, 27 in; in 10 years, 44 in; and in 30 years, 52 in.

Unless it is known that the line will have a life of less than 10 years, not less than 10 years' creep should be allowed for. Creep has come into consideration in transmission-line design only during the past 35 years, and to date no standards have been established for handling it. Probably the simplest approach is to check all close clearance points on the profile with a template made with no creep allowance and to specify higher structures at these points if the addition of liberal creep sag infringes on the required clearances. It is possible to prestress the creep out of small conductors, but for large conductors this requires time and special tensioning facilities not normally available. Also the time lost in constructing an EHV line will more than pay for the extra structure height required to compensate for the creep. Prestressing changes the modulus of elasticity, and this new modulus should be used in the design.

The vertical weight supported at any structure is the weight of the length of conductor between low points of the sag in the two adjacent spans. For bare-conductor weights, this distance between low points can be scaled by using a template of the sag at any desired temperature. The maximum weight under loaded conditions should be scaled from a template made for the loaded sags. For most problems, the horizontal distance may be taken as equal to the conductor length. Distances to the low point of the sag may be computed by Eq. (14-65).

Uplift. On steep inclined spans the low point may fall beyond the lower support; this indicates that the conductor in the uphill span exerts a negative or upward pull on the lower tower. The amount of this upward pull is equal to the weight of the conductor from the lower tower to the low point in the sag. Should the upward pull of the uphill span be greater than the downward load of the next adjacent span, actual uplift would be caused, and the conductor would tend to swing clear of the tower.

It is important that abrupt changes in elevation of the structures should not occur, so that the conductor will not tend to swing clear of any structure even at low temperatures. This condition would be indicated if the 0°F curve of the template can be adjusted to hang free of the center support and just touch the adjacent supports on either side. In northern states it would be well to add a curve to the template for the below-zero temperatures experienced.

Insulator Swing. The uplift condition should not even be approached in laying out suspension insulator construction; that is, each tower should carry a considerable weight of conductor. The minimum weight that should be allowed on any structure may be logically determined by finding the transverse angle to which the insulator string may swing without reducing the clearance from the conductor to the structure too greatly. Also, the ratio of vertical weight to horizontal wind load should be limited to avoid insulator swing beyond this angle. The maximum wind is usually assumed at a temperature of 60°F. The wind pressure, measured in pounds per square foot, to be used in swing calculations is a matter of judgment and depends on local conditions. Under high-wind conditions it is reasonable to require somewhat less than normal clearances. Generally a clearance corresponding to about 75% of the flashover value of the insulator is adequate. The insulator will swing in the direction of the resultant of the vertical and horizontal forces acting on the insulator string as shown in Fig. 14-16.

Long Spans. Rough country may necessitate spans considerably longer than contemplated in the design and may involve a number of factors including (1) proper clearance between conductors, (2) excessive tensions under maximum load, and (3) structures adequate to carry the additional loads.

Safe horizontal clearance between conductors is often based on the National Electrical Safety Code (NESC) formula, in which the spacing a in inches is given as proportional to the square root of sag; s is in inches.

$$a = 0.3\ \text{in/kV} + 8\sqrt{\frac{s}{12}} \tag{14-49}$$

This relation was developed for, and is useful on, comparatively short span lines of the smaller conductors and for voltages up to 69 kV; but for very long spans and heavy conductors, the formula results in spacings considerably larger than have proved satisfactory. It also results in spacings that are questionably small for very light conductors on long spans. Percy H. Thomas proposed an empirical formula which takes into account the weight of the conductor and its diameter, requiring less spacing for heavy conductors and a greater spacing for small conductors by the ratio of diameter D in inches to weight w in pounds per foot (D/w) as a means of determining the required conductor

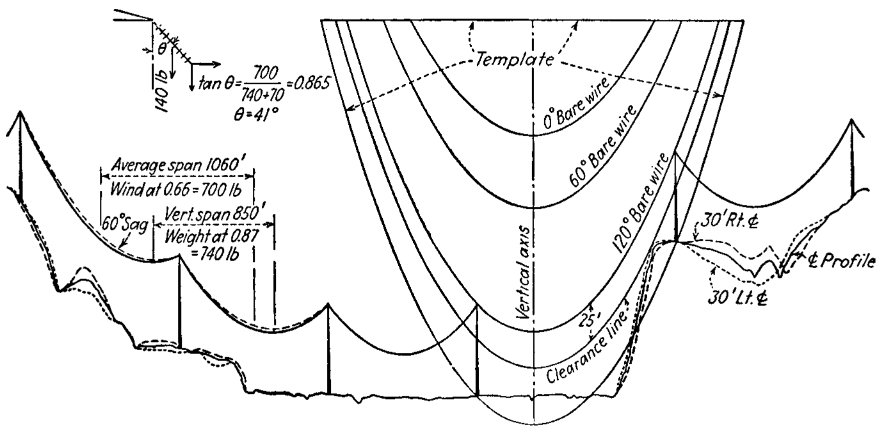


FIGURE 14-16 Sag template determines clearances of a suspended conductor from the ground.

spacing for the average span of the line. The factor C in Eq. (14-81) includes an allowance to permit the standard spacing to be used on somewhat longer spans than average construction. The same formula, however, may be used to examine the spacings which have been successfully used on maximum spans and a value for C selected from experience for determining the safe spacing required for an occasional unusually long span.

Excessive tensions on very long spans may be avoided by dead-ending at both ends and computing such a stringing sag as will result in the same maximum tension as elsewhere in the line. Such a span will be found to have considerably greater stringing sag and lower stringing tension than the normal span. Sag curves or charts are often prepared giving the sag for dead-end spans of various lengths such that the maximum tension under loaded conditions will be the same.

Dead-end construction is costly, and consideration should be given to avoiding this additional expense. It is common practice to permit spans up to double the average span without dead ends, although spans of this length may require additional spacing between wires. A careful examination of some trial figures on the sags and tensions developed in a long span will often indicate how great a span may be carried on suspension structures. The maximum loaded tension which would occur in a long span, if this span were dead-ended and sagged to the same stringing tension as the rest of the line, compared with the maximum tension for normal span lengths, is a good indication of the necessity for dead-end construction.

In case a number of long spans are encountered in a line or section of line, it may prove more economical to reduce the tension in the entire section to the long-span values and accept an increase in sag and corresponding reduction in span length in order to avoid dead ends.

Computerized Tower Spotting.[54–56] In a line of any significant length there are a very large number of possible tower location sequences which meet the requirement for minimum electrical clearances yet also meet the maximum load limits of the chosen structure family. With considerable design experience, it is possible to select a reasonably economical tower spotting solution, but no manual tower spotting method can explore all the possibilities nor find the lowest-cost solution.

In recent years, computer programs have become available to explore nearly all possible tower spotting solutions, selecting those having the lowest cost. In addition to exploring minimum-cost tower spotting solutions for new lines, these computer programs also allow the user to explore uprating alternatives including reconductoring, adding structures, raising attachment points, and retensioning the existing conductors. With the advent of more and more powerful personal computers and easier-to-use graphical interfaces, these programs can be applied even to relatively small line designs. Such programs are particularly attractive when modern digital methods of obtaining terrain data or existing line structure locations, heights, and catenaries can be used to collect the vast amount of input data required.

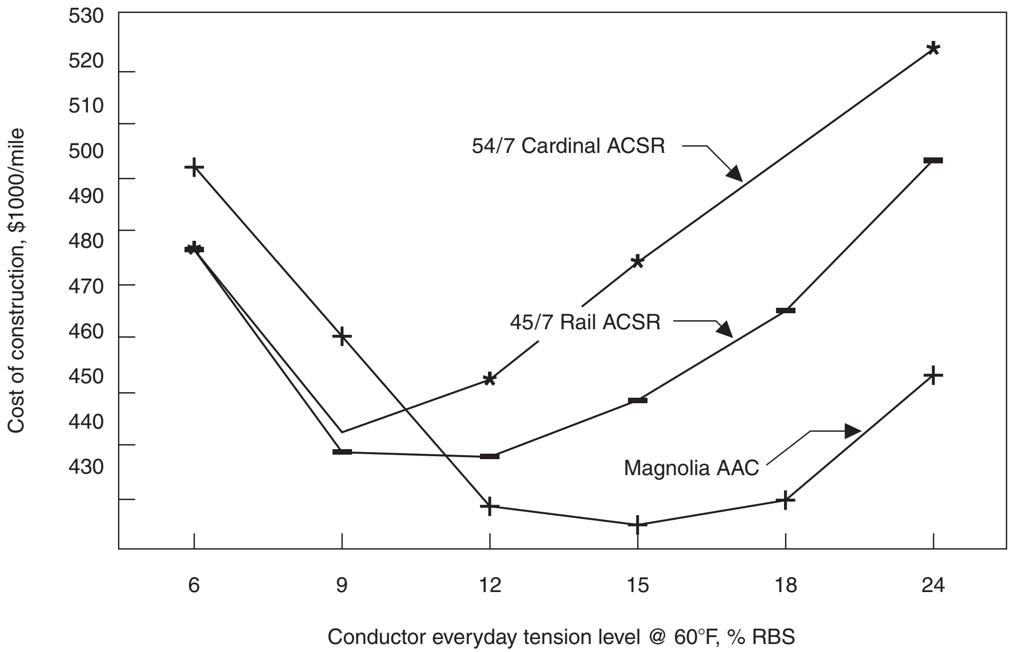


FIGURE 14-17 Cost of construction versus conductor tensions for 1200-ft (366 m) wind span.

Digital data collection and analysis allows the line designer to explore a number of design aspects that were simply impossible just a short time ago. For example, Fig. 14-17 shows the result of a series of lowest-cost numerical tower spotting calculations made to explore the effects of conductor type (all-aluminum conductor, low-steel 45/7 ACSR, and high-steel 54/7 ACSR) and conductor stringing tension expressed as a percent of rated breaking strength (RBS). Each data point represents a optimized tower spotting calculation. It's interesting to note that the lowest-cost solution is the weakest conductor at a modest tension level.

14.1.8 Mechanical Design of Overhead Spans

Conductor and Structure Loads. The span design consists of determining the sag at which the conductor shall be erected so that heavy winds, accumulations of ice or snow, and low temperatures, even if sustained for several days, will not stress the conductor beyond the elastic limit, cause a serious permanent stretch, or result in fatigue failures from continued vibrations.

Unit wind and ice loadings for conductors are found by the following formulas:

$$\text{Wind load (lb/ft)} = p \times \left(\frac{D}{12}\right) \tag{14-50}$$

$$\text{Ice load (lb/ft)} = 1.244 \times (Dr + r^2) \tag{14-51}$$

where p is the wind pressure in pounds per square foot, D is the diameter of the conductor in inches, and r is the radial thickness of the ice in inches. The ice is assumed to be glaze ice with a unit weight of 57 lb/ft^3.

The dead weight of the conductor and the weight of the accumulated ice act vertically; the wind load is assumed to act horizontally and at right angles to the span; the resultant is the vectorial sum.

Under combined vertical and horizontal loading, the conductor swings out into an inclined plane whose angle with the vertical is the angle between the direction of the vertical force and the resultant force. The resulting deflection is measured in this inclined plane.

The following procedures for calculating extreme loadings on transmission line conductors and structures are based on a reliability-based design (RBD) methodology described in ASCE Manual 74.[57] These represent the minimum loading levels for which transmission lines in the United States should be designed. For critical or important lines, more stringent requirements than those given below should be specified to provide improved reliability of the lines. Detailed procedures for designing for higher levels of line reliability are given in Manual 74.

Extreme Wind Loading. The wind pressure p at height z above ground level, in pounds per square foot, is given by the following formula:[57]

$$p = 0.00256(Z_v V)^2 G C_f \tag{14-52}$$

where V = the basic wind speed, in miles per hour, determined from the wind-speed contour map in Fig. 14-18
C_f = the force coefficient given in Table 14-11 or 14-12
Z_v = the terrain factor given in Table 14-13
G = the gust response factor given in Fig. 14-19*a* through *d*

The exposure categories required for the determination of p_z are defined in Table 14-14. These exposure categories and the basic wind-speed map in Fig. 14-18 are not applicable to sections of transmission lines that cross high mountain ridges, large river valleys, or other topographic features where localized wind speed-up effects may occur. In these cases, special meteorological studies should be conducted to establish the appropriate wind loadings.

The basic wind-speed contour map in Fig. 14-18 is taken from ASCE Standard 7-88.[58] Wind speeds from this map represent the 50-year return period fastest-mile speeds at 33 ft above ground for exposure category C.

The effective height z for determining the terrain factor and gust response factor is the distance above ground level to the center of pressure of the conductor or structure. For conductors, it can be approximated as the average height above ground of the conductor attachment points to the structure minus one-third the sum of the insulator length (for suspension insulators only) and the sag of the conductors. For support structures with total heights of 200 ft or less, the effective height can be approximated as two-thirds the total height of the structure. For structures taller than 200 ft, the terrain factor should be varied over the height of the structure to represent the increase in the wind speed with height above ground.

Extreme Ice Loading. The radial ice thickness r for the extreme ice loading condition can be determined from the ice map in Fig. 14-20.[57] This map gives estimates of the average 50-year return period glaze ice thicknesses for five regions of the United States. Since this map was developed from limited observations of icing on overhead lines, it should be used only if statistical data on extreme ice loadings for the region of the transmission line are not available.

Combined Ice and Wind Loading. For combined ice and wind loading conditions, the glaze ice thickness determined from Fig. 14-20 should be combined with a wind speed equal to 0.4 times the wind speed from the wind contour map in Fig. 14-18. The basis for this reduced wind speed is described in ASCE Manual 74.[57] In cases where statistical data on wind speeds during icing conditions are available, those data should be used in lieu of this wind-speed reduction.

Wire Tensions. The wire tensions for the extreme wind loading case should be based on the temperature that is most likely to occur at the time of the extreme wind events. For example, it could be the average of the minimum daily temperatures

TABLE 14-11 Force Coefficients for Cylindrical Surfaces

Description of surface	C_f
Stranded cables (conductors, ground wires, guy wires)	1.0
Smooth circular cylinder	0.9
Rough circular cylinder	1.2
16-sided polygon	0.9
12-sided polygon	1.0
Octagon	1.4

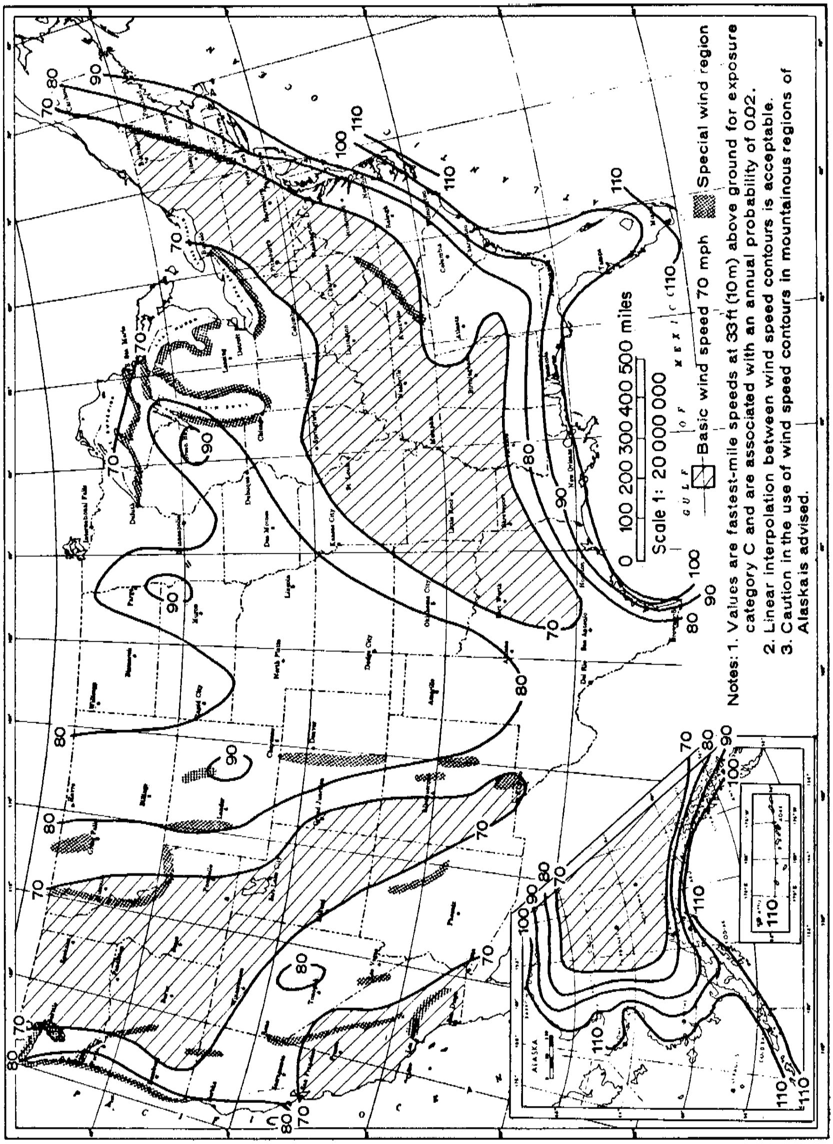


FIGURE 14-18 Basic wind speed, mi/h.

TABLE 14-12 Force Coefficients for Lattice Towers, C_f

ϵ	C_f	
	Square towers	Triangular towers
<0.025	4.0	3.6
0.025−0.44	$4.1-5.2\epsilon$	$3.7-4.5\epsilon$
0.45−0.69	1.8	1.7
0.7−1.0	$1.3+0.7\epsilon$	$1.0+\epsilon$

Notes: ϵ is the ratio of solid area to gross area of tower face.

Force coefficients are given for towers with structural angles or similar flat-sided members.

For towers with rounded members, the design wind force shall be determined using the values in this table multiplied by the following factors:

$\epsilon \leq 0.29$ factor = 0.67
$0.3 \leq \epsilon \leq 0.79$ factor = $0.67\epsilon + 0.47$
$0.8 \leq \epsilon \leq 1.0$ factor = 1.0

For triangular-section towers, the design wind forces shall be assumed to act normal to a tower face.

For square-section towers, the design wind forces shall be assumed to act normal to a tower face. To allow for the maximum horizontal wind load, which occurs when the wind is oblique to the faces, the wind load acting normal to a tower face shall be multiplied by the factor $1.0 + 0.75\epsilon$ for $\epsilon < 0.5$ and shall be assumed to act along a diagonal.

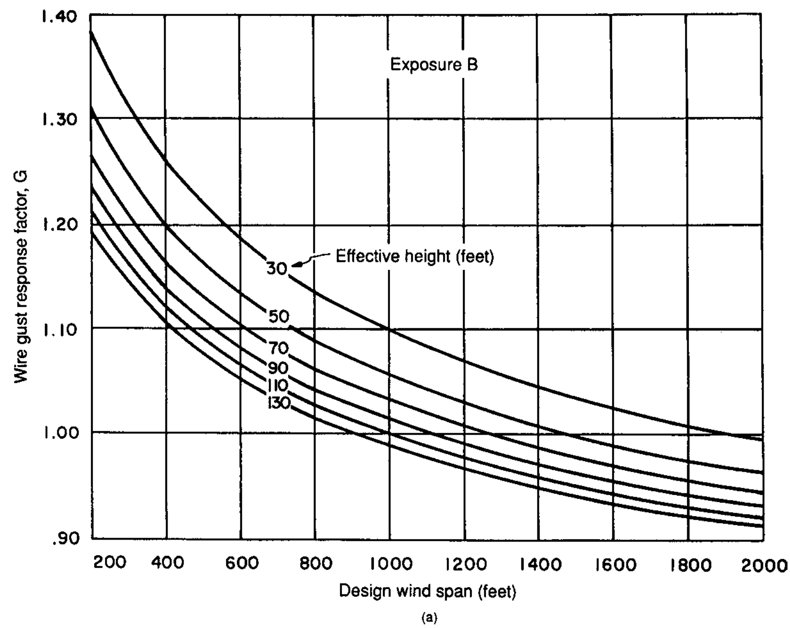
for the strong-wind season. A wire temperature of 15°F is recommended for computing the wire tensions for the combined ice and wind loading case. Although ice accretion typically occurs at temperatures somewhat greater than this, the 15°F temperature accounts for a possible cold front passing after the icing event.

Catenary Calculations for Stranded Conductors. The energized conductors of transmission and distribution lines must be placed in a manner that totally eliminates the possibility of injury to people. Overhead conductors, however, elongate with time, temperature, and tension, thereby changing their original positions after installation. Despite the effects of weather and loading on a line, the conductors must remain at safe distances from buildings, objects, and people or vehicles passing

TABLE 14-13 Terrain Factor, Z_v

Height above ground level, z ft	Z_v		
	Exposure B	Exposure C	Exposure D
0–33	0.72	1.00	1.18
40	0.75	1.03	1.21
50	0.78	1.06	1.23
60	0.82	1.09	1.26
70	0.85	1.11	1.28
80	0.88	1.14	1.29
90	0.91	1.16	1.31
100	0.93	1.17	1.32
120	0.96	1.20	1.35
140	0.99	1.23	1.37
160	1.02	1.26	1.39
180	1.05	1.28	1.40
200	1.08	1.30	1.42

Notes: Linear interpolation for intermediate values of height z is acceptable. Exposure categories are defined in Table 14-1.



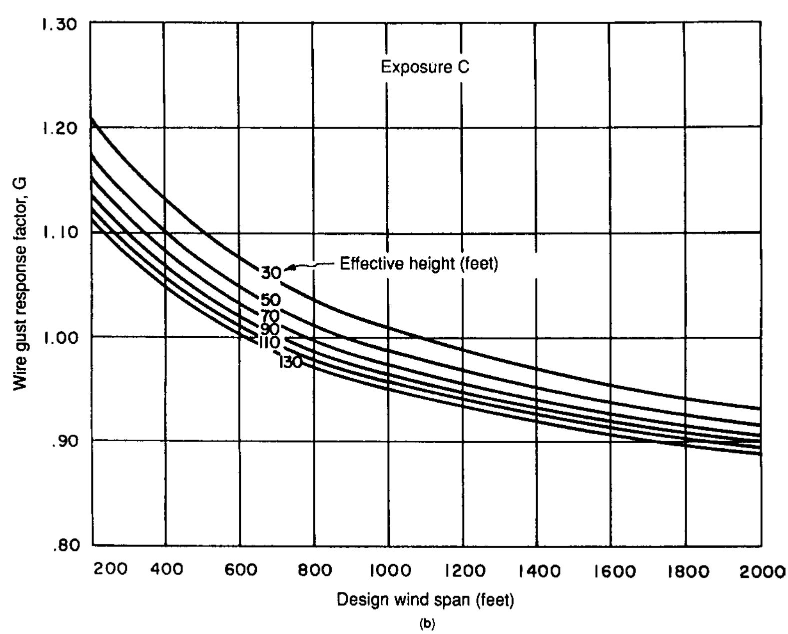
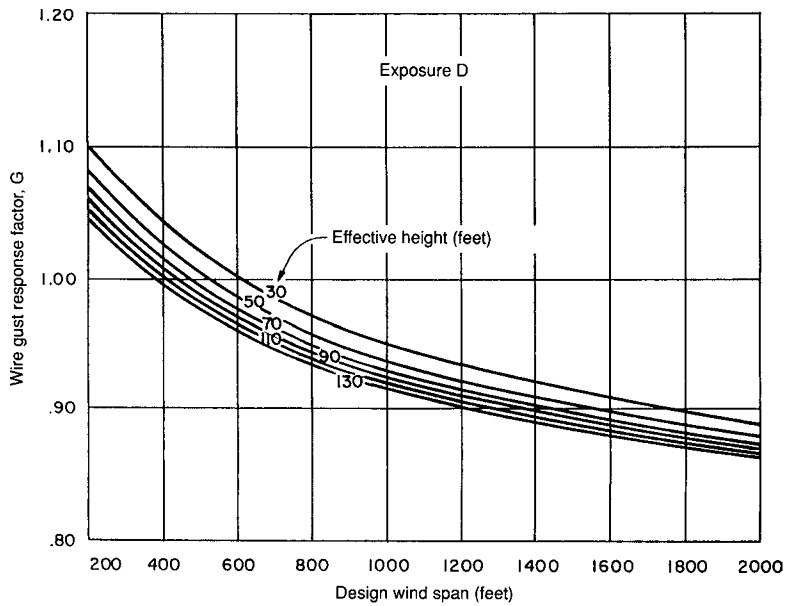
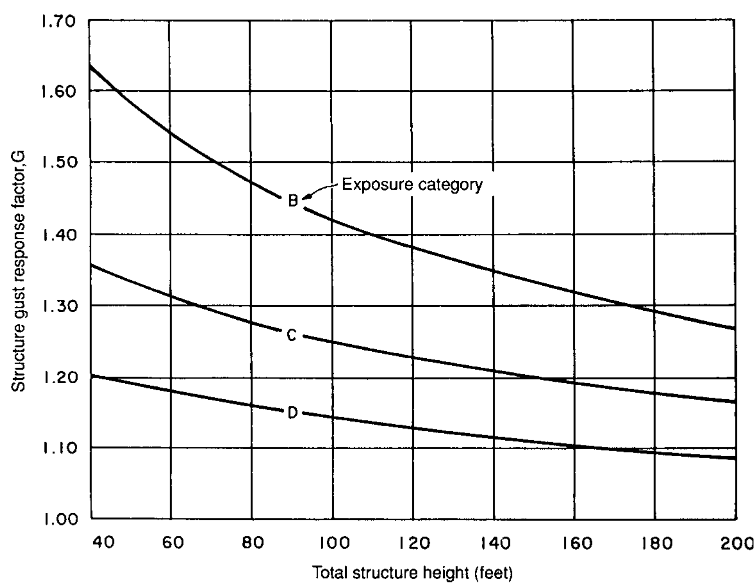


FIGURE 14-19 Conductor gust response factor, exposures B (*a*), C (*b*), and D (*c*); structure gust response factor (*d*).



(c)



(d)

FIGURE 14-19 (*Continued*)

TABLE 14-14 Description of Exposure Categories

Exposure category	Description
B	Suburban areas, wooded areas, or other terrain with numerous closely spaced obstructions having the size of single-family dwellings or larger
C	Open terrain with scattered obstructions having heights generally less than 30 ft, e.g., cultivated fields and grasslands
D	Flat, unobstructed coastal areas directly exposed to wind flowing over large bodies of water

beneath the line at all times. To ensure this safety, the shape of the terrain along the right-of-way, the height and lateral position of the conductor between support points, and the position of the conductor between support points under all wind, ice, and temperature conditions must be known.

Bare overhead transmission or distribution conductors are typically flexible and uniform in weight along their lengths. Because of these characteristics, they take the form of a catenary between support points. The shape of the catenary[59,60] changes with conductor temperature, ice and wind loading, and time. To ensure adequate vertical and horizontal clearance under all weather and electrical loadings, and to ensure that the breaking strength of the conductor is not exceeded, the behavior of the conductor catenary under all conditions must be incorporated into the line design. The required prediction of the future behavior of the conductor are determined through calculations commonly referred to as *sag-tension calculations,* which predict the behavior of conductors according to recommended tension limits under varying loading conditions. These tension limits specify certain percentages of the conductor's rated breaking strength that is not to be exceeded on installation or during the life of the line. These conditions, along with the elastic and permanent elongation properties of the conductor, provide the basis for determining the amount of resulting sag during installation and long-term operation of the line.

Accurately determined initial sag limits are essential in the line design process. Final sags and tensions depend on initial installed sags and tensions and on proper handling during installation. The final sag shape of conductors is used to select support point heights and span lengths so that the minimum clearances will be maintained over the life of the line. If the conductor is damaged or the initial sags are incorrect, the line clearances may be violated or the conductor may break during heavy ice or wind loadings.

Sag and Tension in Level Spans. A bare stranded overhead conductor is normally held clear of objects, people, and other conductors by periodic attachment to insulators. The elevation differences between the supporting structures affect the shape of the conductor catenary. The catenary's shape has a distinct effect on the sag and tension of the conductor, which can be determined using well-defined mathematical equations.

The shape of a catenary is a function of the conductor weight per unit length w, the horizontal component of tension, H, the span length S, and the sag of the conductor D. Conductor sag and span length are illustrated in Fig. 14-21 for a level span.

The exact catenary equation uses hyperbolic functions. Relative to the low point of the catenary curve shown in Fig. 14-21, the height of the conductor $y(x)$ above this low point is given by the following equation:

$$y(x) = \frac{H}{w}\left[\cosh\left(\frac{wx}{H}\right) - 1\right] \cong \frac{wx^2}{2H} \tag{14-53}$$

Note that x is positive in either direction from the low point of the catenary. The expression to the right is an approximate parabolic equation based on a MacLaurin series expansion of the hyperbolic cosine.

For a level span, the low point is in the center and the sag D is found by substituting $x = S/2$ in the preceding equations. The exact catenary and approximate parabolic equations for sag become the following:

$$D = \frac{H}{w}\left[\cosh\left(\frac{wS}{2H}\right) - 1\right] \cong \frac{wS^2}{8H} \tag{14-54}$$

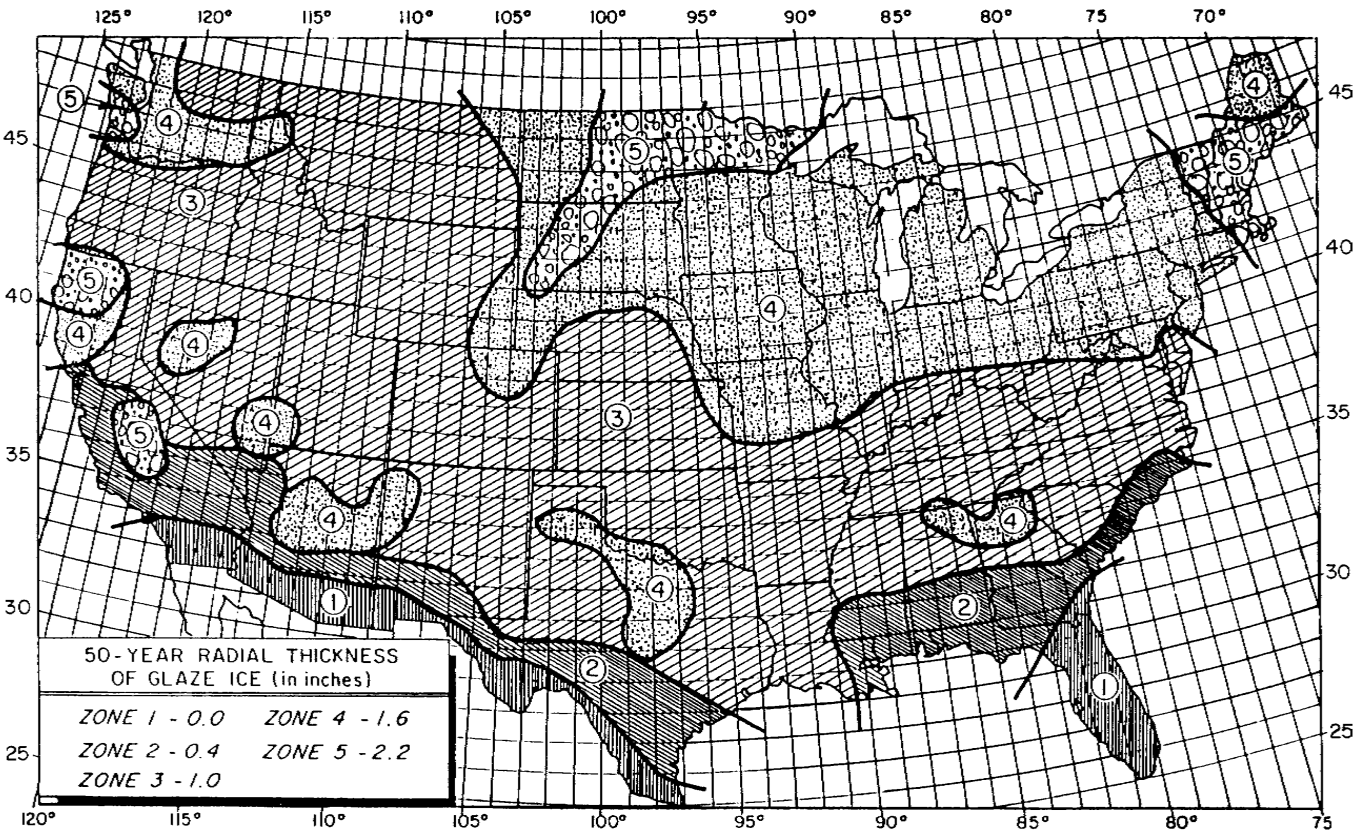


FIGURE 14-20 Extreme radial thickness of glaze ice having a 50-year return period.

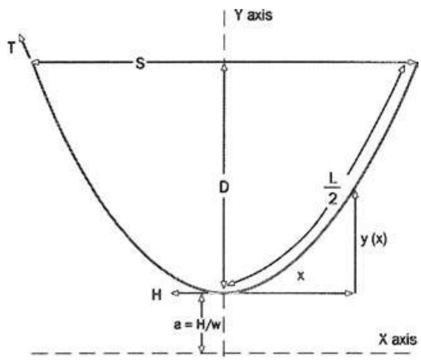


FIGURE 14-21 The catenary curve for level spans.

The ratio H/w which appears in all of the preceding equations is commonly referred to as the *catenary constant*. An increase in the catenary constant causes the catenary curve to become shallower and the sag to decrease. Although it varies with conductor temperature, ice and wind loading, and time, the catenary constant typically has a value in the range of several thousand feet for most transmission-line catenaries.

The approximate, or parabolic, expression is sufficiently accurate as long as the sag is less than 5% of the span length. As an example, consider a 1000-ft (304.5-m) span of Drake ACSR conductor with a per unit weight of 1.096 lb/ft (15.99 N/m) installed at a tension of 4500 lb (20.016 kN). The catenary constant H/w is 4106 ft (1251.8 m). The calculated sag is 30.48 ft (9.293 m) and 30.44 ft (9.280 m) using the hyperbolic and approximate parabolic equations, respectively. For this case where the sag-to-span ratio is 3.4%, the difference in calculated sag between the hyperbolic and parabolic equations is 0.48 in (1.3 cm).

The horizontal component of tension H is equal to the conductor tension at the point in the catenary where the conductor slope is horizontal. For a level span, this is the midpoint of the span. At the ends of the level span, the conductor tension T is equal to the horizontal component plus the conductor weight per unit length w multiplied by the sag D, as shown in the following:

$$T = H + wD \tag{14-55}$$

Given the conditions in the preceding example calculation for a 1000-ft (304.8-m) level span of ACSR Drake, the tension at the attachment points T exceeds the 4500-lb (20.016-N) horizontal component of tension H by only 36 lb (162 N), a difference of only 0.8%.

This shows that the use of horizontal tension H and parabolic equations for the catenary are adequate for typical transmission spans and sags. However, there is little reason to use either approximation in numerical methods.

Conductor Length. Application of calculus to the catenary equation allows the calculation of the conductor length $L(x)$ measured along the conductor from the low point of the catenary in either direction.

The equation for catenary length between the supports is

$$L(x) = \frac{H}{w}\sinh\left(\frac{wx}{H}\right) \cong x\left(1 + \frac{x^2w^2}{6H^2}\right) \tag{14-56}$$

For a level span, the conductor length corresponding to $x = S/2$, is half of the total conductor length L; thus

$$L = \left(\frac{2H}{w}\right)\sinh\left(\frac{Sw}{2H}\right) \cong S\left(1 + \frac{S^2w^2}{24H^2}\right) \tag{14-57}$$

The parabolic equation for conductor length can also be expressed as a function of sag D by substitution of the sag parabolic equation [Eq. (14-54)]:

$$L = S + \frac{8D^2}{3S} \tag{14-58}$$

Conductor slack is the difference between the conductor length L and the span length S. The parabolic equations for slack may be found by equating and rearranging the preceding parabolic equations for conductor length L and sag D:

$$L - S \cong S^3\left(\frac{w^2}{24H^2}\right) \cong D^2\left(\frac{8}{3S}\right) \tag{14-59}$$

While slack has units of length, it may also be expressed as a percentage of the span length. In the preceding catenary calculation, the length of the catenary is 1002.471 ft (305.63 m) and the slack is 2.471 ft (0.753 m), which is 2.47% of the span length. According to the ruling-span approximation, which is discussed later, the tension H and the tension to weight per unit length ratio H/w is the same in all suspension spans between strain structures. Therefore the slack in each suspension span is proportional to the cube of the suspension span length and the total slack is determined largely by the longest spans. It is for this reason that the ruling span is closer to the longest span rather than the average span.

Equation (14-58) can be inverted to obtain an equation showing the dependence of sag D on slack $L - S$:

$$D = \sqrt{\frac{3S(L-S)}{8}} \tag{14-60}$$

Given the preceding 1000-ft (304.5-m) level span of Drake ACSR conductor with 2.471 ft (0.753 m) of slack, a reduction of 6 in (15.2 cm) in slack yields a sag reduction of 3.25 ft (0.99 m). As can be seen from the preceding, small changes in slack typically yield large changes in conductor sag and tension, particularly for short spans.

Sag and tension in inclined spans may be analyzed using essentially the same equations that were used for level spans. The catenary equation for the conductor height above the low point in the span is the same. However, the span is considered to consist of two separate sections, one to the right of the low point and the other to the left as shown in Fig. 14-22. The shape of the catenary relative to the low point is unaffected by the difference in suspension point elevation (span inclination). In each direction from the low point, the conductor elevation $y(x)$ relative to the low point is given by Eq. (14-53):

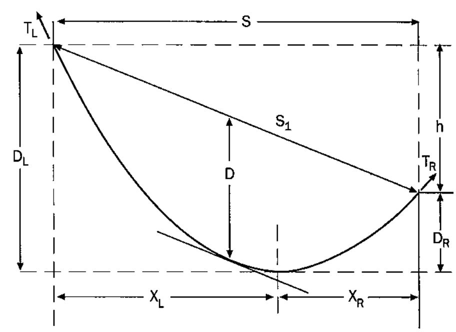


FIGURE 14-22 Inclined catenary span.

$$y(x) = \frac{H}{w}\left[\cosh\left(\frac{wx}{H}\right) - 1\right] = \frac{wx^2}{2H}$$

Note that x is considered positive in either direction from the low point.

The horizontal distance x_L from the left support point to the low point in the catenary is

$$x_L = \frac{S}{2}\left(1 + \frac{h}{4D}\right) \tag{14-61}$$

The horizontal distance x_R from the right support point to the low point of the catenary is

$$x_R = \frac{S}{2}\left(1 - \frac{h}{4D}\right) \tag{14-62}$$

where S = horizontal distance between support points
h = vertical distance between support points
S_1 = straight-line distance between support points
D = sag measured vertically from a line through the points of conductor support to a line tangent to the conductor

(as shown in Fig. 14-22). The midpoint sag D is approximately equal to the sag in a horizontal span, with a length equal to the inclined span S_1.

Knowing the horizontal distance from the low point to the support point in each direction, we can apply the preceding equations for $y(x)$, L, D, and T to each side of the inclined span. The total conductor length L in the inclined span is equal to the sum of the lengths in the x_R and x_L subspan sections:

$$L = S + (x_L^3 + x_R^3)\left(\frac{w^2}{6H^2}\right) \tag{14-63}$$

In each subspan, the sag is relative to the corresponding support point elevation

$$D_R = \frac{Wx_R^2}{2H}, \qquad D_L = \frac{Wx_L^2}{2H} \tag{14-64}$$

or in terms of sag D and the vertical distance between support points

$$D_R = D\left(1 - \frac{h}{4D}\right)^2, \qquad D_L = D\left(1 + \frac{h}{4D}\right)^2 \tag{14-65}$$

and the maximum tension is

$$T_R = H + wD_R, \qquad T_L = H + wD_L \tag{14-66}$$

or in terms of upper and lower support points:

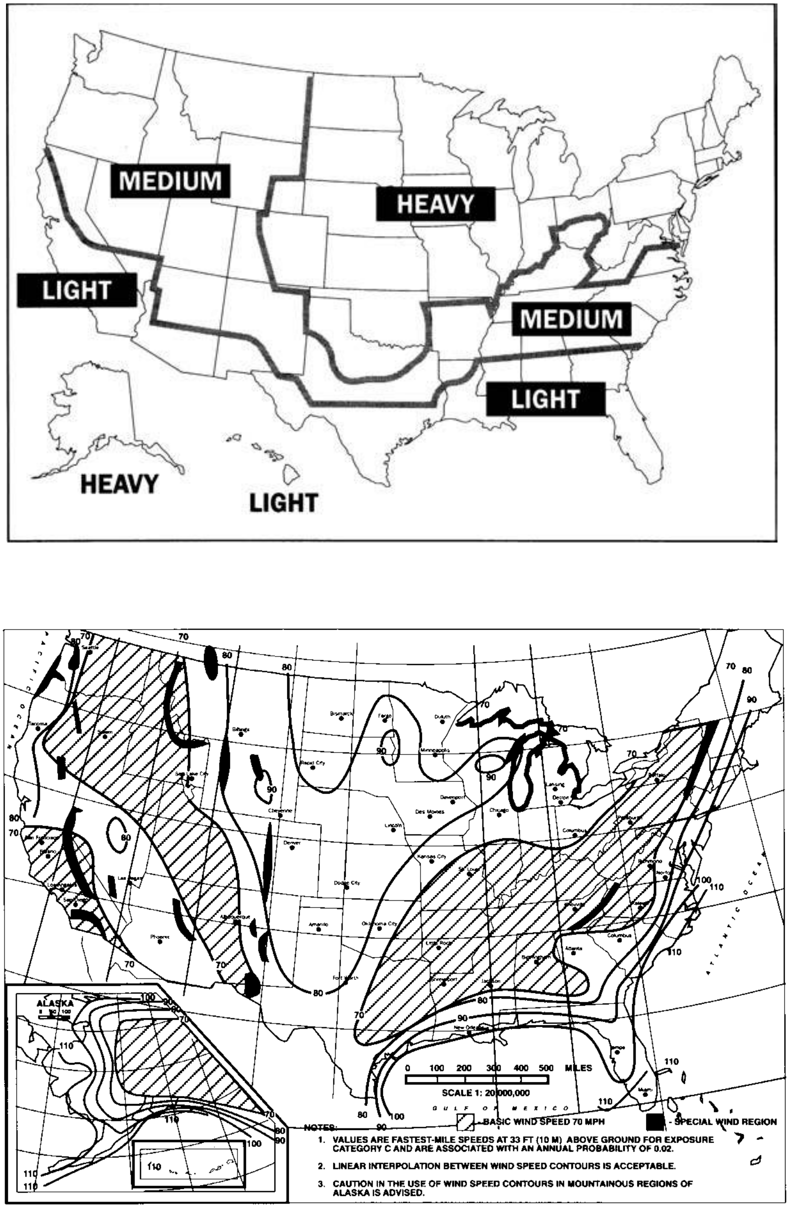
$$T_u = T_1 + wh \tag{14-67}$$

where D_R = sag in right subspan section
D_L = sag in left subspan section
T_R = tension in right subspan section
T_L = tension in left subspan section
T_u = tension in conductor at upper support
T_1 = tension in conductor at lower support

The horizontal conductor tension is equal at both supports. The vertical component of conductor tension is greater at the upper support, and the resultant tension, T_u, is also greater.

Ice and Wind Conductor Loads. When a conductor is covered with ice and/or is exposed to wind, the effective conductor weight per unit length increases. During occasions of heavy ice and/or wind load, the conductor catenary tension increases dramatically along with the loads on angle and dead-end structures. Both the conductor and its supports can fail unless these high-tension conditions are considered in the line design.

The *National Electrical Safety Code* (NESC) suggests certain combinations of ice and wind corresponding to heavy, medium, and light loading regions of the United States. Figure 14-23*a* is a map of the United States indicating those areas. The combinations of ice and wind corresponding to loading region are listed in Table 14-15. The NESC also suggests that increased conductor loads due to high wind loads without ice be considered. Figure 14-23*b* shows the suggested wind pressure as a function of geographic area for the U.S. (NESC).



This figure is reproduced by permission of the American Society of Civil Engineers. [10]

FIGURE 14-23 (*a*) Ice and wind load areas of the United States. (*b*) Wind-pressure design values for the United States. (*American Society of Civil Engineers*).

Certain utilities in very heavy ice areas use glaze ice thickness of as much as 2 in (50 mm) to calculate iced conductor weight. Similarly, utilities in regions where hurricane winds occur may use wind loads as high as 34 lb/ft^2 (1620 Pa). As the NESC indicates, the degree of ice and wind loads varies by region. Some areas may have heavy icing, whereas some areas may have extremely high winds. The loads must be accounted for in the line design process to prevent a detrimental effect on

TABLE 14-15 Ice and Wind Loading For NESC Loading Districts

	Loading districts			
	Heavy	Medium	Light	Extreme wind loading
Radial thickness of ice, in (mm)	0.5 (12.5)	0.25 (6.25)	0	0
Horizontal wind pressure, lb/ft^2 (Pa)	4 (190)	4 (190)	9 (430)	see Fig. 2-4 (NESC)
Temperature, °F (°C)	0 (−20)	+15 (−10)	+30 (−1)	+60 (+15)
Constant to be added to the resultant (all conductors), lb/ft (N/m)	0.30 (4.40)	0.20 (2.50)	0.05 (0.70)	0.0 (0.0)

the line. Some of the effects of both the individual and combined components of ice and wind loads are discussed below:

Ice loading of overhead conductors may take several physical forms (glaze ice, rime ice, or wet snow). The impact of lower-density ice formation is usually considered in the design of line sections at high altitudes.

The formation of ice on overhead conductors has the following influence on line design:

- Ice loads determine the maximum vertical conductor loads that structures and foundations must withstand.
- In combination with simultaneous wind loads, ice loads also determine the maximum transverse loads on structures.
- In regions of heavy ice loads, the maximum sags and the permanent increase in sag with time (difference between initial and final sags) may be due to ice loadings.

Ice loads for use in designing lines are normally derived on the basis of past experience, code requirements, state regulations, and analysis of historical weather data. Mean recurrence intervals for heavy ice loadings are a function of local conditions along various routings. The impact of varying assumptions concerning ice loading can be investigated with line design software.

The calculation of ice loads on conductors is normally done with an assumed glaze ice density of 57 lb/ft^3 (8950 N/m^3). The weight of ice per unit length is calculated with the following equation:

$$w_{\text{ice}} = 0.0281t(D_c + t) \qquad (14\text{-}68)$$

where t = radial thickness of ice, mm
D_c = conductor outside diameter, mm
w_{ice} = resultant weight of ice, N/m

The ratio of iced weight to bare weight depends strongly on conductor diameter. As shown in Table 14-16, for three different conductors covered with 0.5-in radial glaze ice, this ratio ranges from 4.8 for No. 1/0 AWG wire to 1.6 for 1590-kcmil conductors. As a result, small-diameter conductors may need to have a higher elastic modulus and higher tensile strength than do large conductors in heavy ice and wind loading areas to limit sag.

TABLE 14-16 Ratio of Iced to Bare Conductor Weight

ACSR Conductor	D_c, in	W_{bare}, lb/ft	W_{ice}, lb/ft	$\frac{W_{\text{bare}} + W_{\text{ice}}}{W_{\text{bare}}}$
No. 1/0 AWG 6/1 "Raven"	0.398	0.1452	0.559	4.8
477-kcmil 26/7 "Hawk"	0.858	0.6570	0.845	2.3
1590-kcmil 54/19 "Falcon"	1.545	2.044	1.272	1.6

Conductor wind loading influences line design in a number of ways:

- The maximum span between structures may be determined by the need for horizontal clearance to edge of right-of-way during moderate winds.
- The maximum transverse loads for tangent and small-angle suspension structures are often determined by infrequent high-wind-speed loadings.
- Permanent increase in conductor sag may be determined by wind loading in areas of light ice load.

Wind-pressure load on conductors P_w is commonly specified in pounds per square foot (lb/ft^2). The relationship between P_w and wind velocity is given by the following equation:

$$P_w = 0.00256(V_w)^2 \tag{14-69}$$

where V_w is the wind speed in miles per hour.

The wind load per unit length of conductor W_w is equal to the wind-pressure load P_w multiplied by the conductor diameter (including radial ice of thickness t, if any):

$$W_w = P_w \frac{(D_c + 2t)}{12} \tag{14-70}$$

Combined Ice and Wind Loading. If the conductor weight is to include both ice and wind loading, the resultant magnitude of the loads must be determined. The weight of a conductor under both ice and wind loading is given by the following equation:

$$w_{w+i} = \sqrt{(w_b + w_i)^2 + (W_w)^2} \tag{14-71}$$

where w_b = bare conductor weight per unit length, lb/ft (N/m)
w_i = weight of ice per unit length, lb/ft (N/m)
w_w = wind load per unit length, lb/ft (N/m)
w_{w+i} = resultant of ice and wind loads, lb/ft (N/m)

The NESC prescribes a safety factor K in pounds per foot, dependent on loading district, to be added to the resultant ice and wind loading when performing sag and tension calculations. Therefore, the total resultant conductor weight w is

$$w = w_{w+i} + K \tag{14-72}$$

Conductor Tension Limits. The NESC recommends limits on the tension of bare overhead conductor as a percentage of the conductor's rated breaking strength. The tension limits are 60% under maximum ice and wind loading, 35% initial unloaded (when installed) at 60°F, and 25% final unloaded (after maximum loading has occurred), also at 60°F. It is common, however, for lower unloaded tension limits to be used. Except in areas experiencing severe ice loading, it is not unusual to find tension limits of 60% maximum, 25% unloaded initial, and 15% unloaded final. This set of specifications could easily result in an actual maximum tension on the order of only 35% to 40%, an initial tension of 20%, and a final unloaded tension level of 15%. In this case, the 15% tension limit is said to govern.

Transmission-line conductors are seldom covered with ice, and winds on the conductor are usually much lower than those used in maximum load calculations. Under such everyday conditions, tension limits are specified to limit eolian vibration to safe levels. Even with everyday lower tension levels of 15% to 20%, it is assumed that vibration control devices will be used in those sections of the line which are subject to severe vibration. Eolian vibration levels, and thus appropriate unloaded tension limits, vary with the type of conductor, the terrain, the span length, and the use of dampers. Special conductors such as ACSS, SDC, and VR, which exhibit high self-damping properties, may be installed to the full code limits, if desired.

Recent studies by CIGRE Working Group B2/11 led to the development of a guide for choosing suitable installed conductor tension levels. The recommendations are in the form of maximum allowable ratios of tension (*H* in pounds or Newtons) to weight per unit length (*w* in lb/ft or N/m). The primary determinants of the maximum *H/w* ratio include the type of conductor (e.g., ACSR or AAC). The type of terrain (e.g., 4 terrain types, #1 being open, flat, not trees, no obstructions, etc.), and a span parameter *LD/m* where *L* is the span length, *D* is the conductor diameter, and *m* is the mass per unit length of the conductor.

Approximate Sag-Tension Calculations. Sag-tension calculations, using detailed experimental stress-strain and creep elongation laboratory data, are usually performed with the aid of a computer; however, with certain simplifications, these calculations can be made with a handheld calculator. The latter approach allows greater insight into the calculation of sags and tensions than is possible with complex computer programs. Equations suitable for such calculations, as presented in the preceding section, can be applied to the following example:

It is desired to calculate the sag and slack for a 600-ft level span of 795 kcmil-26/7 ACSR Drake conductor. The bare conductor weight per unit length w_b is 1.094 lb/ft. The conductor is installed with a horizontal tension component *H* of 6300 lb, equal to 20% of its rated breaking strength of 31,500 lb. The sag for this level span is

$$D = \frac{1.094(600)^2}{(8)6300} = 7.81 \text{ ft } (2.38 \text{ m})$$

The length of the conductor between the support points is determined from

$$L = 600 + \frac{8(7.81)^2}{3(600)} = 600.27 \text{ ft } (182.96 \text{ m})$$

Note that the conductor length depends solely and directly on span and sag. It is not directly dependent on conductor tension, weight, or temperature. The *conductor slack* is the conductor length minus the span length; in this example, it is 0.27 ft (0.082 m).

Sag Change with Thermal Elongation. The ACSR and AAC conductors elongate with increasing conductor temperature. The rate of linear thermal expansion for the composite ACSR conductor is less than that of the AAC conductor because the steel strands in the ACSR elongate at approximately half the rate of aluminum. The composite coefficient of linear thermal expansion of a nonhomogenous conductor, such as ACSR Drake, may be found from the following equations:[61]

$$\alpha_{AS} = \alpha_{Al}\left(\frac{F_{Al}}{E_{AS}}\right)\left(\frac{A_{Al}}{A_{total}}\right) + \alpha_{St}\left(\frac{E_{St}}{E_{AS}}\right)\left(\frac{A_{St}}{A_{total}}\right) \tag{14-73}$$

$$E_{AS} = E_{Al}\left(\frac{A_{Al}}{A_{total}}\right) + E_{St}\left(\frac{A_{St}}{A_{total}}\right) \tag{14-74}$$

where E_{Al} = modulus of elasticity of aluminum, lb/in^2
E_{St} = modulus of elasticity of steel, lb/in^2
E_{AS} = modulus of elasticity of aluminum-steel composite, lb/in^2
A_{Al} = area of aluminum strands, square units
A_{St} = area of steel strands, square units
A_{total} = total cross-sectional area, square units
α_{Al} = aluminum coefficient of linear thermal expansion, °F^{-1}
α_{St} = steel coefficient of thermal elongation, °F^{-1}
α_{AS} = composite aluminum-steel coefficient of thermal elongation, °F^{-1}

Of course, the modulus of elasticity of an ACSR conductor is not linear. Its elongation is a complex function of both the present stress and prior stress loading over time. Nonetheless, it is informative

to calculate the thermal elongation for approximate elasticity moduli for steel and aluminum strands.
Using elastic moduli of 8 and 28 million lb/in^2 for aluminum and steel, respectively, the elastic modulus for ACSR Drake is

$$E_{AS} = (8.6 \times 10^6)\left(\frac{0.6247}{0.7264}\right) + (27 \times 10^6)\left(\frac{0.1017}{0.7264}\right) = 11.2 \times 10^6 \text{ lb/in}^2$$

and the coefficient of linear thermal expansion is

$$\alpha_{AS} = 12.8 \times 10^{-6}\left(\frac{8.6 \times 10^6}{11.2 \times 10^6}\right)\left(\frac{0.6247}{0.7264}\right) + 6.4 \times 10^{-6}\left(\frac{27 \times 10^6}{11.2 \times 10^6}\right)\left(\frac{0.1017}{0.7264}\right)$$

$$= 10.6 \times 10^{-6}\,°\text{F}^{-1}$$

If the conductor temperature changes from a reference temperature T_{ref} to another temperature T, the conductor length L changes in proportion to the product of the conductor's effective coefficient of linear thermal elongation α_{AS} and the change in temperature $T - T_{ref}$ as follows:

$$L_T = L_{T,ref}[1 + \alpha_{AS}(T - T_{ref})] \tag{14-75}$$

For example, if the temperature of the Drake conductor in the preceding example increases from 60°F (15°C) to 167°F (75°C), then the length increases by 0.68 ft (0.21 m) from 600.27 ft (182.96 m) to 600.95 ft (183.17 m):

$$L(\text{at } 167°\text{F}) = 600.27[1 + (10.6 \times 10^{-6})(167 - 60)] = 600.95 \text{ ft}$$

Ignoring for the moment any change in length due to change in tension, the sag at 167°F (75°C) may be calculated for the conductor length of 600.95 ft (183.17 m) using Eq. (14-60):

$$D = \sqrt{\frac{3(600)(0.95)}{8}} = 14.6 \text{ ft (4.45 m)}$$

Using a rearrangement of Eq. (14-54), this increased sag is found to correspond to a decreased tension of

$$H = \frac{w(S^2)}{8D} = \frac{1.094(600)^2}{8(14.6)} = 3372 \text{ lb (15,100 N)}$$

If the conductor were inextensible, that is, if it had an infinite modulus of elasticity, then these values of sag and tension for a conductor temperature of 167°F would be correct. For any real conductor, however, the elastic modulus of the conductor is finite and changes in tension change the conductor length. Use of the preceding calculation, therefore, will overstate the increase in sag.

Sag Change Due to Combined Thermal and Elastic Effects. With moduli of elasticity around the 10-million-lb/in^2 level, typical bare aluminum and ACSR conductors elongate about 0.01% for every 1000-lb/in^2 change in tension. In the preceding example, the increase in temperature caused an increase in length and sag and a decrease in tension, but the effect of tension change on length was ignored. As discussed later, concentric-lay stranded conductors, particularly nonhomogenous conductors such as ACSR, are not inextensible. Rather, they exhibit quite complex elastic and plastic behavior. Initial loading of conductors results in elongation behavior substantially different from that caused by loading many years later. Also, high-tension levels caused by heavy ice and wind loads cause a permanent increase in conductor length, affecting subsequent elongation under various conditions.

Accounting for such complex stress-strain behavior usually requires a sophisticated, computer-aided approach. For illustration purposes, however, the effect of permanent elongation of the conductor on sag and tension calculations will be ignored, and a simplified elastic conductor assumed. This idealized conductor is assumed to elongate linearly with load and to undergo no permanent increase in length, regardless of loading or temperature. For such a conductor, the relationship between tension and length is

$$L_H = L_{H,\text{ref}}\left[1 + \frac{H - H_{\text{ref}}}{E_C A}\right] \tag{14-76}$$

where L_H = length of conductor under horizontal tension, H
$L_{H,\text{ref}}$ = length of conductor under horizontal reference tension, H_{ref}
E_c = modulus of elasticity of the conductor, lb/in^2
A = cross-sectional area, in^2

In calculating sag and tension for extensible conductors, it is useful to add a step to the preceding calculation of sag and tension for elevated temperature. This added step allows a separation of thermal elongation and elastic elongation effects, and involves the calculation of a zero-tension length (ZTL) at the conductor temperature of interest T_{cdr}. This ZTL (T_{cdr}) is the conductor length attained if the conductor is taken down from its supports and laid on the ground with no tension. By reducing the initial tension in the conductor to zero, the elastic elongation is also reduced to zero, shortening the conductor. It is even possible that for short spans, the zero tension length can be less than the span length.

Consider the preceding example for ACSR Drake in a 600-ft level span. The initial conductor temperature is 60°F, the conductor length is 600.27 ft, and E_{AS} is calculated to be 11.2 × 10^6 lb/in^2. Using Eq. (14-76), the reduction of the initial tension from 6300 lb to zero yields a ZTL at 60°F of

$$\text{ZTL}_{(60°\text{F})} = 600.27\left[1 + \frac{0 - 6300}{(11.2 \times 10^6)\,(0.7264)}\right] = 599.80 \text{ ft (182.82 m)}$$

Keeping the tension zero and increasing the conductor temperature to 167°F yields a purely thermal elongation. The zero tension length at 167°F can be calculated using Eq. (14-75):

$$\text{ZTL}_{(167°\text{F})} = 599.80 \text{ ft } [1 + (10.6 \times 10^{-6})\,(167 - 60)] = 600.48 \text{ ft}$$

According to Eqs. (14-54) and (14-60), this length corresponds to a sag of 10.2 ft and a horizontal tension of 4412 lb. However, this length was calculated for zero tension; it will elongate elastically under tension. The actual conductor sag-tension determination requires a process of iteration as follows:

1. As described above, the conductor's ZTL, calculated at 167°F (75°C), is 600.48 ft; sag is 10.2 ft; and the horizontal tension H is 4412 lb.
2. Because the conductor is elastic, the tension of 4412 lb will increase the conductor length from 600.48 ft to

$$\text{L}_{1(167°\text{F})} = 600.48\left[1 + \frac{4412 - 0}{(0.7264)(11.2 \times 10^6)}\right] = 600.80 \text{ ft (183.12 m)}$$

3. The sag D_1(167°F) corresponding to this length is calculated using Eq. (14-60):

$$\text{D}_{1(167°\text{F})} = \sqrt{\frac{3(600)(0.80)}{8}} = 13.4 \text{ ft (4.09 m)}$$

4. Using Eq. (14-54), this sag yields a new horizontal tension H_1(167°F) of

$$H_1 = \frac{1.094(600)^2}{8(13.4)} = 3674 \text{ lb}$$

TABLE 14-17 Iterative Solution for Increased Conductor Temperature

Iteration no.	Length L_n, ft	Sag D_n, ft	Tension H_n, lb	New trial tension, lb
ZTL	600.48	10.2	4412	—
1	600.80	13.4	3674	$\frac{4412 + 3674}{2} = 4043$
2	600.77	13.2	3730	$\frac{3730 + 4043}{2} = 3887$
3	600.76	13.1	3758	$\frac{3758 + 3887}{2} = 3823$
4	600.75	13.0	3787	$\frac{3787 + 3823}{2} = 3805$

A new trial tension is taken as the average of H and H_1, and the process is repeated. The results are described in Table 14-17.

Note that the balance of thermal and elastic elongation of the conductor yields an equilibrium tension of approximately 3800 lb and a sag of 13.0 ft. The calculations of the previous section, which ignored elastic effects, result in lower tension (3372 lb) and a greater sag (14.6 ft).

Slack is equal to the excess of conductor length over span length. Table 14-17 can be replaced by a plot of the catenary and elastic curves on a graph of slack versus tension. The solution occurs at the intersection of the two curves. Figure 14-24 shows the tension-versus-slack curves intersecting at a tension of 3800 lb, which agrees with the preceding calculations.

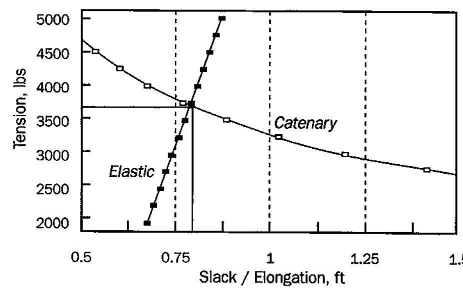


FIGURE 14-24 Sag-tension solution for 600-ft (183-m)

Sag Change Due to Ice Loading. As a final example of sag-tension calculation, calculate the sag and tension for the 600-ft span of Drake with the addition of 0.5 in of radial ice and a drop in a conductor temperature to 0°F. The weight of the conductor increases by

$$w_{\text{ice}} = 1.244t\,(D_c + t) = 1.244(0.5)(1.108 + 0.5)$$
$$= 1.000 \text{ lb/ft } (14.6 \text{ N/m})$$

As in the previous example, the calculation uses the conductor's ZTL at 60°F, which is the same as that found in the previous section, 599.80 ft. The ice loading is specified for a conductor temperature of 0°F, so the ZTL at 0°F, using Eq. (14-75), is

$$\text{ZTL}_{(0°\text{F})} = 599.80[1 + (10.6 \times 10^{-6})(0 - 60)]$$
$$= 599.42 \text{ ft}$$

As in the case of sag and tension at elevated temperatures, the conductor tension is a function of slack and elastic elongation. The conductor tension and the conductor length are found at the point of intersection of the catenary and elastic curves (Fig. 14-25). The intersection of the curves occurs at a horizontal tension component of 12,275 lb, not

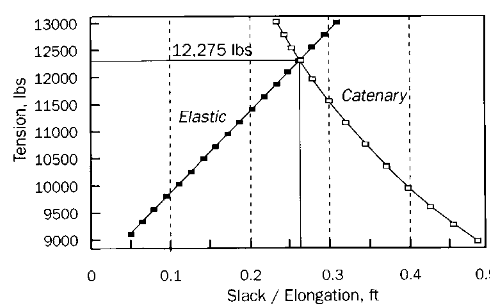


FIGURE 14-25 Sag-tension solution for 600-ft (183-m) span of Drake ACSR at 0°F (−17.8°C) and 0.5-in ice.

very far from the crude initial estimate of 12,050 lb, that ignored elastic effects. The sag corresponding to this tension and the iced conductor weight per unit length is 9.2 ft.

In spite of doubling the conductor weight per unit length by adding 0.5 in of ice, the sag of the conductor is much less than the sag at 167°F. This condition is generally true for transmission conductors where minimum ground clearance is determined by the high temperature rather than the heavy loading condition. Small distribution conductors, such as the 1/0–AWG ACSR in Table14-16, experience a much larger ice-to-conductor weight ratio (4:8), and the conductor sag under maximum wind and ice load may exceed the sag at moderately higher temperatures.

The preceding approximate tension calculations could have been more accurate with the use of actual stress-strain curves and graphic sag-tension solutions, as described in detail in Ref. 62. This method, although accurate, is very slow and has been replaced completely by computational methods.[63]

Experimental Stress-Strain Curves. Sag-tension calculations are normally done numerically and allow the user to enter many different loading and conductor temperature conditions. Both initial and final conditions are calculated, and multiple tension constraints can be specified. The complex stress-strain behavior of ACSR-type conductors can be modeled numerically, including both temperature and elastic and plastic effects.

Stress-strain curves for bare overhead conductors include a minimum of an initial curve and a final curve over a range of elongations from 0% to 0.45%. For conductors consisting of two materials, an initial and final curve for each is included. Creep curves for various lengths of time are typically included as well.

Overhead conductors are not purely elastic. They stretch with tension, but when the tension is reduced to zero, they do not return to their initial length. Thus, conductors are plastic; the change in conductor length cannot be expressed with a simple linear equation, as used in the preceding hand calculations. The permanent length increase that occurs in overhead conductors yields the difference in initial and final sag-tension data found in most computer programs.

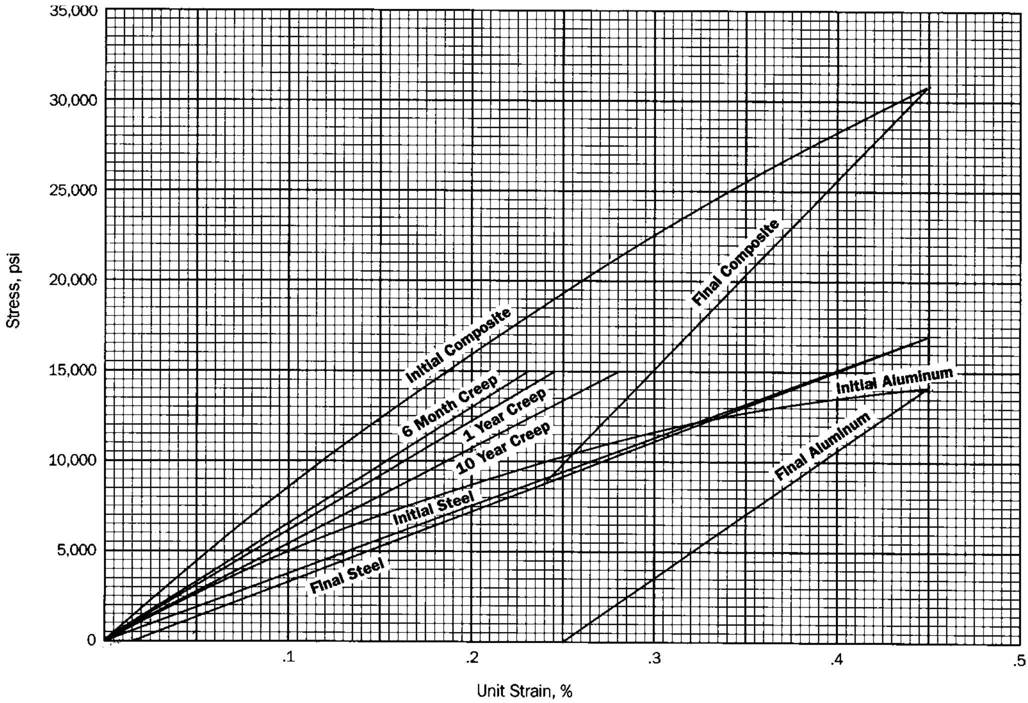
Figure 14-26 shows a typical stress-strain curve[64] for a 26/7 ACSR conductor; the curve is valid for conductor sizes ranging from 266.8 to 795 kcmil. A 795-kcmil 26/7 ACSR Drake conductor has a breaking strength of 31,500 lb (14,000 kg) and an area of 0.7264 in^2 (46.9 mm) so that it fails at an average stress of 43,000 lb/in^2 (30 kg/mm^2). The stress-strain curve illustrates that at a stress equal to 50% of the conductor's breaking strength (21,500 lb/in^2), the elongation is less than 0.3%. This translates to an elongation of 1.8 ft (0.55 m) in a 600-ft (180-m) span.

Note that the component curves for the steel core and the aluminum-stranded outer layers are separated. This separation allows for changes in the relative curve locations as the temperature of the conductor changes.

For the preceding example, with the Drake conductor at a tension of 6300 lb (2860 kg), the length of the conductor in the 600-ft (180-m) span was found to be 0.27 ft longer than the span. This tension corresponds to a stress of 8600 lb/in^2 (6.05 kg/mm^2). From the stress-strain curve in Fig. 14-26, this corresponds to an initial elongation of 0.105% (0.63 ft). As in the preceding hand calculation, if the conductor tension is zero, its unstressed length would be less than the span length.

Figure 14-27 is a stress-strain curve[64] for an all-aluminum 37-strand conductor ranging in size from 250 to 1033.5 kcmil. Because the conductor is made entirely of aluminum, there is only one initial curve and one final curve.

Permanent Conductor Elongation Due to High Tensions. Once a conductor has been installed to an initial tension, it can elongate further. Such elongation results from two phenomena: permanent elongation due to high-tension levels resulting from ice and wind loads and creep elongation under everyday tension levels. These types of conductor elongation are discussed in the following sections. Both Figs. 14-26 and 14-27 indicate that when the conductor is initially installed, it elongates nonlinearly. If the conductor tension increases to a relatively high level under ice and wind loading, the conductor will elongate. When the wind and ice loads abate, the conductor elongation will reduce along a curve parallel to the final curve, but will never return to its original length.



Equations for Curves (X = unit strain in %; Y = stress in psi):

Initial Composite: $X = 4.07 \times 10^{-3} + (1.28 \times 10^{-5})\,Y - (1.18 \times 10^{-10})\,Y^2 + (5.64 \times 10^{-15})\,Y^3$

$Y = -512 + (8.617 \times 10^4)\,X - (1.18 \times 10^4)\,X^2 - (5.76 \times 10^{-4})\,X^3$

Initial Steel: $Y = (37.15 \times 10^3)\,X$

Initial Aluminum: $Y = -512 = (4.902 \times 10^4)\,X - (1.18 \times 10^4)\,X^2 - (5.76 \times 10^4)\,X^3$

Final Composite: $Y = (107.55X - 17.65) \times 10^3$

Final Steel: $Y = (38.60X - 0.65) \times 10^3$

Final Aluminum: $Y = (68.95X - 17.00) \times 10^3$

6 Month Creep: $Y = (64.75 \times 10^3)\,X$

1 Year Creep: $Y = (60.60 \times 10^3)\,X$

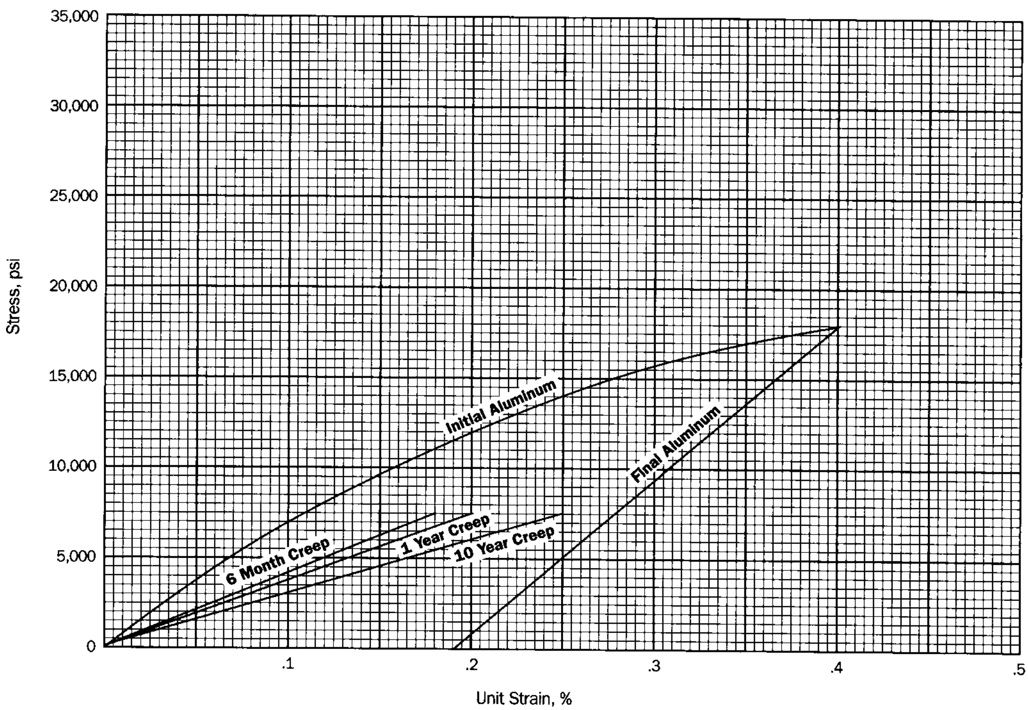
10 Year Creep: $Y = (53.45 \times 10^3)\,X$

Test Temperature 70°F to 75° F

FIGURE 14-26 Stress-strain curves for 26/7-strand ACSR.

For example, refer to Fig. 14-27 and assume that a newly strung 795-kcmil, 37-strand Arbutus AAC has an everyday tension of 2780 lb. The conductor area is 0.6245 in^2, so the everyday stress is 4450 lb/in^2 and the elongation is 0.062%. Following an extremely heavy ice and wind load event, assume that the conductor stress reaches 18000 lb/in^2. When the conductor tension decreases back to everyday levels, the conductor elongation will be permanently increased by more than 0.2%. Also the sag under everyday conditions will be correspondingly greater, and the tension will be less. In most numerical sag-tension methods, final sag-tension values are calculated for such permanent elongation due to heavy loading conditions.

Permanent Elongation at Everyday Tensions (Creep). Conductors permanently elongate under tension even if the tension level never exceeds everyday levels. This permanent elongation caused by everyday tension levels is called *creep*.[64] Creep can be determined by long-term laboratory creep tests.



Equations for Curves (X = unit strain in %; Y = stress in psi):

Initial Aluminum: $X = -5.31 \times 10^{-3} + (1.74 \times 10^{-5})\,Y - (6.17 \times 10^{-10})\,Y^2 + (5.05 \times 10^{-14})\,Y^3$

$Y = 136 + (7.46 \times 10^{4})\,X - (8.51 \times 10^{4})\,X^2 + (2.33 \times 10^{4})\,X^3$

Final Aluminum: $Y = (85.20\,X - 16.14) \times 10^{3}$

6 Month Creep: $Y = (42.30 \times 10^{3})\,X$

1 Year Creep: $Y = (38.20 \times 10^{3})\,X$

10 Year Creep: $Y = (30.60 \times 10^{3})\,X$

Test Temperature 70°F to 75° F

FIGURE 14-27 Stress-strain curves for 37-strand AAC.

The results of the tests are used to generate creep-versus-time curves. On the stress-strain graphs, creep curves are often shown for 6-month, 1-year, and 10-year periods. Figure 14-27 shows these typical creep curves for a 37-strand 250- to 1033.5-kcmil AAC. In Fig. 14-27, assume that the conductor tension remains constant at the initial stress of 4450 lb/in^2. At the intersection of this stress level and the initial elongation curve, 6-month, 1-year, and 10-year creep curves, the conductor elongation from the initial elongation of 0.062% increases to 0.11%, 0.12%, and 0.15%, respectively. Because of creep elongation, the resulting final sags are greater and the conductor tension is less than the initial values.

Creep elongation in aluminum conductors is quite predictable as a function of time and obeys a simple exponential relationship. Thus, the permanent elongation due to creep at everyday tension can be found for any period of time after initial installation. Creep elongation of copper and steel strands is much less and is normally ignored. Permanent increase in conductor length due to heavy load occurrences cannot be predicted at the time a line is built. The reason for this unpredictability is that the occurrence of heavy ice and wind loads is random. A heavy ice storm may occur the day after the line is built or may never occur over the life of the line.

Sag-Tension Tables. To illustrate the result of typical sag-tension calculations, Tables 14-18 to 14-21 present initial and final sag-tension data for 795-kcmil 26/7 ACSR Drake, 795-kcmil, 37-strand AAC

Arbutus, and 795-kcmil Type 16 SDC Drake conductors in NESC light and heavy loading areas for spans of 1000 and 300 ft. Typical tension constraints of 15% final unloaded at 60°F, 25% initial unloaded at 60°F, and 60% initial at maximum loading are used. The calculations in these tables were performed with the Alcoa SAG10™ program, version 1.1.

Initial versus Final Sags and Tensions. Rather than calculated as a function of time, most sag-tension calculations are based on initial and final loading conditions. Initial sags and tensions are simply the sags and tensions at the time the line is built. Final sags and tensions are calculated assuming that (1) the specified ice and wind loading has occurred and (2) the conductor has experienced 10 years of creep elongation at a conductor temperature of 60°F at the user-specified initial tension.

With most sag-tension calculation methods, final sags are calculated for both heavy ice/wind loads and for creep elongation. The final sag-tension values reported to the user are those with the greatest increase in sag.

ACSR Sag-Tension Calculations. Sag-tension calculations for ACSR conductors are more complex than those for AAC, AAAC, or ACAR conductors. The complexity results from the different behavior of steel and aluminum strands in response to tension and temperature. Steel wires exhibit neither creep elongation nor plastic elongation in response to high tensions. Aluminum wires do creep and respond plastically to high stress levels. Also, they elongate twice as much as steel wires do in response to changes in temperature. Table 14-18 presents various initial and final sag-tension values for a 600-ft span of an ACSR Drake conductor under heavy loading conditions. Note that the tensions in the aluminum and steel components are given separately. In particular, some other useful observations are

1. At 60°F, without ice or wind, the tension level in the aluminum strands decreases with time as the strands permanently elongate as a result of creep or heavy loading.
2. Both initially and finally, the tension level in the aluminum strands decreases with increasing temperature (212°F), reaching zero tension at 167°F for initial and final conditions, respectively.
3. At the highest temperature (212°F), where all the tension is in the steel core, the initial and final sag and tension values are nearly the same, illustrating the fact that the steel core does not permanently elongate in response to time or high tension.

TABLE 14-18 Sag and Tension Data for 795 kcmil 26/7 ACSR Drake Conductor
Span = 600 ft; NESC heavy loading district; creep is not a factor.*

					Final		Initial	
Temperature, °F	Ice, in	Wind, lb/ft^2	*k*, lb/ft	Resultant weight, lb/ft	Sag, ft	Tension, lb	Sag, ft	Tension, lb
0	0.50	4.00	0.30	2.509	11.14	10153	11.14	10153
32	0.50	0.00	0.00	2.094	44.54	8185	11.09	8512
−20	0.00	0.00	0.00	1.094	6.68	7372	6.27	7855
0	0.00	0.00	0.00	1.094	7.56	6517	6.89	7147
30	0.00	0.00	0.00	1.094	8.98	5490	7.95	6197
60	0.00	0.00	0.00	1.094	10.44	4725†	9.12	5402
90	0.00	0.00	0.00	1.094	11.87	4157	10.36	4759
120	0.00	0.00	0.00	1.094	13.24	3727	11.61	4248
167	0.00	0.00	0.00	1.094	14.29	3456	13.53	3649
212	0.00	0.00	0.00	1.094	15.24	3241	15.24	3241

* See Appendix for more detailed tables.
† Design condition.

Sag change of ACSR at high temperature

When first installed, the aluminum and steel wire components of ACSR are both under tension and elongate equally. Ignoring for the moment any initial creep of the aluminum strands, the tension division between aluminum and steel in Drake ACSR when initially installed at 20% RBS (6300 lb) at 60°F may be calculated.

If the elastic modulus of the steel core is taken as 27 Mpsi and the steel core area is 0.1017 in^2, then the "spring constant" of the steel core is 2.75 Mlb. If the elastic modulus of the aluminum layers is taken as 8.6 Mpsi and the aluminum strand area is 0.6247 in^2, then the spring constant of the aluminum layers is 5.37 Mpsi. Given the two "springs" in parallel, the tension division can be calculated as follows:

$$H_A = \frac{6300}{\left(1 + \dfrac{2.75}{5.37}\right)} = 4165 \text{ lb}$$

And the tension in the steel core is 2135 lb.

If the temperature of this conductor is increased from 60 to 212°F, the unstressed length of the steel core would increase to 600.38 ft = 599.80 [1 + 6.4e − 6(212 − 60)] and the aluminum strand layers would increase to 600.97 ft = 599.80[1 + 12.8e − 6(212 − 60)]. Clearly the unstressed lengths of the steel core and aluminum layers are now different. Reapplying tension to the composite condutor, the steel core must be preloaded to 2598 lb = (0.59ft/600.38)(26e6 × 0.1017 in^2) before the tension in the aluminum layers beings to increase. The net result of this behavior is that, at high temperature, the conductor tension in the aluminum strands can become zero. This temperature is commonly called the ACSR conductor "kneepoint" temperature. Beyond this "kneepoint" temperature, the thermal elongation of the ACSR conductor is reduced to a level near to that of the steel core alone, though low levels of compression in the aluminum layers must be considered.

Mechanical Interaction of Suspension Spans

Transmission lines are normally designed in line sections with each end of the section terminated by a strain structure that allows no longitudinal movement of the conductor.[60] Structures within each line section are typically suspension structures that support the conductor vertically but allow free movement of the conductor attachment point either longitudinally or transversely.

Tension Differences for Adjacent Dead-End Spans. Table 14-19 contains initial and final sag and tension data for a 700-ft (213-m) and a 1000-ft (305-m) dead-end span with an ACSR Drake conductor that was initially installed with the same 6300-lb tension limits at 60°F. Note that the differences between final tensions at 60°F is 260 lb, which is due entirely to the difference in span length. Even the initial tension (equal at 60°F) differs by approximately 880 lb at –20°F and 610 lb at 167°F.

Tension Equalization by Suspension Insulators. At a typical suspension structure, the conductor is supported vertically by a suspension insulator assembly, but allowed to move freely in the direction of the conductor axis. This conductor movement is possible because of insulator swing along the conductor axis. Changes in conductor tension between spans, caused by changes in temperature, load, and time are normally equalized by insulator swing, eliminating horizontal tension differences across suspension structures.

Ruling-Span Approximation. The sag and tension for a series of suspension spans in a line section can be found using the ruling-span concept.[59,60] The ruling-span (RS) for the line section is defined by the following equation:

$$\text{RS} = \sqrt{\frac{S_1^3 + S_2^3 + \cdots + S_n^3}{S_1 + S_2 + \cdots + S_n}}$$

TABLE 14-19 Tension Differences in Adjacent Dead-End Spans for 795-kcmil 26/7 ACSR Drake Conductor

Temperature, °F	Ice, in	Wind, lb/ft²	*k*, lb/ft	Resultant weight, lb/ft	Final Sag, ft	Final Tension, lb	Initial Sag, ft	Initial Tension, lb
Span = 700 ft; NESC heavy loading district; area = 0.7264 in²; creep *is* a factor.								
0	0.50	4.00	0.30	2.509	13.61	11318	13.55	11361
32	0.50	0.00	0.00	2.094	13.93	9224	13.33	9643
−20	0.00	0.00	0.00	1.094	8.22	8161	7.60	8824
0	0.00	0.00	0.00	1.094	9.19	7301	8.26	8115
30	0.00	0.00	0.00	1.094	10.75	6242	9.39	7142
60	0.00	0.00	0.00	1.094	12.36	5429	10.65	6300
90	0.00	0.00	0.00	1.094	13.96	4809	11.99	5596
120	0.00	0.00	0.00	1.094	15.52	4330	13.37	5020
167	0.00	0.00	0.00	1.094	16.97	3960	15.53	4326
212	0.00	0.00	0.00	1.094	18.04	3728	17.52	3837
Span = 1000 ft; area = 0.7264 in²; NESC heavy loading district; creep is *not* a factor.								
0	0.50	4.00	0.30	2.509	25.98	12116	25.98	12116
32	0.50	0.00	0.00	2.094	26.30	9990	25.53	10290
−20	0.00	0.00	0.00	1.094	18.72	7318	17.25	7940
0	0.00	0.00	0.00	1.094	20.09	6821	18.34	7469
30	0.00	0.00	0.00	1.094	22.13	6197	20.04	6840
60	0.00	0.00	0.00	1.094	24.11	5689	21.76	6300*
90	0.00	0.00	0.00	1.094	26.04	5271	23.49	5839
120	0.00	0.00	0.00	1.094	30.14	4923	27.82	5444
167	0.00	0.00	0.00	1.094	30.14	4559	27.82	4935
212	0.00	0.00	0.00	1.094	31.47	4369	30.24	4544

*Design condition.

where RS = ruling span for the line section containing n suspension spans
S_1 = span length of first suspension span
S_2 = span length of second suspension span
S_n = span length of nth suspension span

Alternatively, a generally satisfactory method for estimating the ruling span is to take the sum of the average suspension span length plus two-thirds of the difference between the maximum span and the average span. However, some judgment must be exercised in using this method because a large difference between the average and maximum span may cause a substantial error in the ruling span value.

As discussed earlier, suspension spans are supported by suspension insulators that are free to move in the direction of the conductor axis. This freedom of movement allows the tension in each suspension span to be equal to that calculated for the ruling span. This assumption is valid for the suspension spans and ruling span under the same conditions of temperature and load, for both initial and final sags. For level spans, sag in each suspension span is given by the parabolic sag equation

$$D_i = \frac{w(S_i^2)}{8H_{\mathrm{RS}}} \tag{14-77}$$

where D_i = sag in the ith span
S_i = span length of the ith span
H_{RS} = horizontal tension from ruling span sag-tension calculations

Suspension spans vary in length, although typically not over a large range. Conductor temperature during sagging varies over a range considerably smaller than that used for line design purposes.

If the sag in any suspension span exceeds approximately 5% of the span length, a correction factor should be added to the sag obtained from Eq. (14-77), or the sag should be calculated using the catenary method presented in Eq. (14-79). This correction factor may be calculated as follows:

$$\text{Correction} = D^2\left(\frac{w}{8H}\right) \tag{14-78}$$

where D = sag obtained from parabolic equation
w = weight of conductor, lb/ft
H = horizontal tension, lb

The catenary equation for calculating the sag in a suspension or stringing span is

$$\text{Sag} = \frac{H}{w}\left[\cosh\frac{Sw}{2H} - 1\right] \tag{14-79}$$

where S = span length, ft
H = horizontal tension, lb
w = resultant weight, lb/ft

Stringing Sag Tables. Conductors are typically installed in line section lengths consisting of multiple spans. The conductor is pulled from the conductor reel at a point near one strain structure, progressing through travelers attached to each suspension structure to a point near the next strain structure. After stringing, the conductor tension is increased until the sag in one or more suspension spans reaches the appropriate stringing sags according to the ruling span for the line section. The calculation of stringing sags is based on the preceding sag equation.

Table 14-21 shows a typical stringing sag table for a 600-ft ruling span of ACSR Drake with suspension spans ranging from 400 to 700 ft and conductor temperatures of 20 to 100°F. All the values in this stringing table have been calculated using the parabolic sag equation with ruling-span initial tensions shown in Table 14-20.

Line Design Sag-Tension Parameters. In laying out a transmission line, the first step is to survey the route and create a plan-profile of the selected right-of-way. The plan-profile drawings serve an important function in linking together the various stages involved in the design and construction of the line. These drawings, based on the route survey, show the location and elevation of all natural

TABLE 14-20 Sag and Tension Data for 795-kcmil 26/7 ACSR Drake 600-ft Ruling Span
NESC heavy loading district; area = 0.7264 in^2; creep is *not* a factor.

					Final		Initial	
Temperature, °F	Ice, in	Wind, lb/ft^2	k, lb/ft	Resultant weight, lb/ft	Sag, ft	Tension, lb	Sag, ft	Tension, lb
0	0.50	4.00	0.30	2.509	11.14	10153	11.14	10153
32	0.50	0.00	0.00	2.094	44.54	8185	11.09	8512
–20	0.00	0.00	0.00	1.094	6.68	7372	6.27	7855
0	0.00	0.00	0.00	1.094	7.56	6517	6.89	7147
30	0.00	0.00	0.00	1.094	8.98	5490	7.95	6197
60	0.00	0.00	0.00	1.094	10.44	4725	9.12	5402
90	0.00	0.00	0.00	1.094	11.87	4157	10.36	4759
120	0.00	0.00	0.00	1.094	13.24	3727	11.61	4248
167	0.00	0.00	0.00	1.094	14.29	3456	13.53	3649
212	0.00	0.00	0.00	1.094	15.24	3241	15.24	3241

TABLE 14-21 Stringing Sag Table for 795 kcmil-26/7 ACSR Drake 600-ft Ruling Span
Controlling design condition; 15% RBS at 60°F; No ice or wind, final; NESC heavy loading district.

Horizontal tension, lb	6493	6193	5910	5645	5397	5166	4952	4753	4569
Temperature, °F	20	30	40	50	60	70	80	90	100
Spans					Sag, ft-in				
400	3-4	3-6	3-8	3-11	4-1	4-3	4-5	4-7	4-9
410	3-6	3-9	3-11	4-1	4-3	4-5	4-8	4-10	5-0
420	3-9	3-11	4-1	4-3	4-6	4-8	4-10	5-1	5-3
430	3-11	4-1	4-3	4-6	4-8	4-11	5-1	5-4	5-6
440	4-1	4-3	4-6	4-8	4-11	5-2	5-4	5-7	5-10
450	4-3	4-6	4-8	4-11	5-2	5-4	5-7	5-10	6-1
460	4-5	4-8	4-11	5-2	5-4	5-7	5-10	6-1	6-4
470	4-8	4-11	5-1	5-4	5-7	5-10	6-1	6-4	6-7
480	4-10	5-1	5-4	5-7	5-10	6-1	6-4	6-8	6-11
490	5-1	5-4	5-7	5-10	6-1	6-4	6-8	6-11	7-2
500	5-3	5-6	5-9	6-1	6-4	6-7	6-11	7-2	7-6
510	5-6	5-9	6-0	6-4	6-7	6-11	7-2	7-6	7-9
520	5-8	6-0	6-3	6-7	6-10	7-2	7-6	7-9	8-1
530	5-11	6-2	6-6	6-10	7-1	7-5	7-9	8-1	8-5
540	6-2	6-5	6-9	7-1	7-5	7-9	8-1	8-5	8-9
550	6-4	6-8	7-0	7-4	7-8	8-0	8-4	8-8	9-1
560	6-7	6-11	7-3	7-7	7-11	8-4	8-0	9-0	9-5
570	6-10	7-2	7-6	7-10	8-3	8-7	9-0	9-4	9-9
580	7-1	7-5	7-9	8-2	8-6	8-11	9-4	9-8	10-1
590	7-4	7-8	8-1	8-5	8-10	9-3	9-7	10-0	10-5
600	7-7	7-11	8-4	8-9	9-1	9-6	9-11	10-4	10-9
610	7-10	8-3	8-7	9-0	9-5	9-10	10-3	10-9	11-2
620	8-1	8-6	8-11	9-4	9-9	10-2	10-7	11-1	11-6
630	8-4	8-9	9-2	9-7	10-1	10-6	11-0	11-5	11-11
640	8-8	9-1	9-6	9-11	10-5	10-10	11-4	11-9	12-3
650	8-11	9-4	9-9	10-3	10-9	11-2	11-8	12-2	12-8
660	9-2	9-7	10-1	10-7	11-1	11-6	12-0	12-6	13-1
670	9-5	9-11	10-5	10-11	11-5	11-11	12-5	12-11	13-5
680	9-9	10-3	10-8	11-2	11-9	12-3	12-9	13-4	13-10
690	10-0	10-6	11-0	11-6	12-1	12-7	13-2	13-8	14-3
700	10-4	10-10	11-4	11-11	12-5	13-0	13-6	14-1	14-8

and artificial obstacles to be traversed by, or adjacent to, the proposed line. These plan-profiles are drawn to scale and provide the basis for tower spotting and line design work.

Once the plan-profile is completed, one or more estimated ruling spans for the line may be selected. On the basis of these estimated ruling spans and the maximum design tensions, sag-tension data may be calculated providing initial and final sag values. From these data, sag templates may be constructed to the same scale as the plan-profile for each ruling span, and used to graphically spot structures.

Catenary Constants. The sag in a ruling span is equal to the weight per unit length w times the span length S squared, divided by 8 times the horizontal component of the conductor tension H. The ratio of conductor horizontal tension H to weight per unit length w is the catenary constant H/w. For a ruling-span sag-tension calculation using eight loading conditions, a total of 16 catenary values could be defined, one for initial and one for final tensions under each loading condition.

Catenary constants can be defined for each loading condition of interest and are used in locating structures. Some typical uses of catenary constants are to avoid overloading, ensure that ground clearance is sufficient at all points along the right-of-way, and minimize blowout or uplift under cold-weather conditions. To do this, the following catenary constants are typically found: (1) the maximum line temperature; (2) heavy ice and wind loading; (3) wind blowout; and (4) minimum conductor temperature. Under any of these loading conditions, the catenary constant allows sag calculation at any point within the span.

Wind and Weight Span. The maximum wind span of any structure is equal to the distance measured from center to center of the two adjacent spans supported by the structure. The wind span is used to determine the maximum horizontal force a structure must withstand under high-wind conditions. The wind span is not dependent on conductor sag or tension, only on the horizontal span length.

The weight span of a structure is a measure of the maximum vertical force a structure must withstand. The weight span is equal to conductor weight per unit length times the horizontal distance between the low points of sag of the two adjacent spans. The maximum weight span for a structure is dependent on the design ice and wind loading condition. When the elevations of adjacent structures are the same, the wind and weight spans are equal.

Uplift at Suspension Structures. Conductor uplift, shown in Fig. 14-28, occurs when the weight span of a structure is negative. On steeply inclined spans, the low point of sag may fall beyond the lower support. This indicates that the conductor in the uphill span is exerting a negative, or upward, force on the lower tower. The amount of this upward force is equal to the weight of the conductor from the lower tower to the low point in the sag. If the upward pull of the uphill span is greater than the downward load of the next adjacent span, actual uplift will occur and the conductor will swing free of the tower. This usually occurs under minimum temperature conditions and must be dealt with by adding weights to the insulator suspension string or using a strain structure.

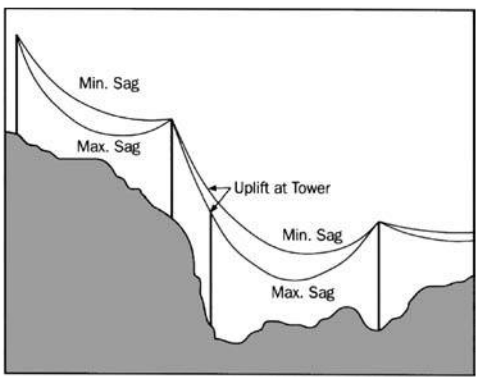


FIGURE 14-28 Illustration of conductor uplift.

Tower Spotting. Given sufficiently detailed plan-profile drawings, structure heights, wind/weight spans, catenary constants, and minimum ground clearances, structure locations can be chosen such that ground clearance is maintained and structure loads are acceptable. Tower spotting can be performed using a sag template, plan-profile drawing, and structure heights, or numerically, by using one of several commercial computer programs. Tower spotting and optimization of line design parameters are discussed in Sec. 14.1.7.

Unbalanced Ice Loads.[65,66] The jump of the conductor resulting from ice dropping off one span of an ice-covered line has been the cause of many serious outages on long-span lines where conductors are arranged in the same vertical plane. The vertical spacing required to prevent "ice-jump" trouble may be estimated by static calculations of the differential sag of two vertically adjacent conductors assuming one conductor has ice and the other has no ice. Normally, sufficient clearance is provided to accommodate this sag difference, including a factor for sag error with a margin for switching surge withstand.

Utility practice, based on historically satisfactory field performance, is typically based on the following criteria for vertical phase-to-phase clearance due to ice:

A maximum sag error of 6 in is assumed.

The upper conductor is assumed to be subject to maximum ice load, typically 1 in for short ruling spans, 0.5 in for unusually long spans.

The lower conductor is assumed to be completely free of ice.

Clearance sufficient to withstand the maximum switching surge is to be provided.

The calculation is performed using normal sag-tension procedures. For certain line designs (e.g., compact lines with post insulators), dynamic techniques are possible.[44,50] However, the trouble has been practically cleared up by horizontally offsetting the conductors from 18 in to 3 ft on medium-voltage lines. The conductor jumps in practically a vertical plane, and this should be true if no wind is blowing, since then all forces and reactions are in a vertical direction.

Wind-Induced Motion of Overhead Conductors. In addition to ordinary "blowout" of overhead conductors (i.e., the swinging motion of the conductor due to the normal component of wind), there are three types of cyclic wind-induced motion that can be a source of damage to structures or conductors or that can result in sufficient reduction in electrical clearance to cause flashover. The categories of wind-induced cyclic motion are eolian vibration, galloping, and wake-induced oscillation.

Eolian vibration can occur when conductors are exposed to a steady low-velocity wind. If the amplitude of such vibration is sufficient, it can result in strand fatigue and/or fatigue of conductor accessories. The amplitude of vibration can be reduced by reducing the conductor tension, adding damping by using dampers (or clamps with damping characteristics), or by the use of special conductors which either provide more damping than standard conductors or are shaped so as to prevent resonance between the tensioned conductor span and the wind-induced vibration force.

Galloping is normally confined to conductors with a coating of glaze ice over at least part of their circumference and thus is not a problem in those areas where ice storms do not occur. It may be controlled by the use of various accessories attached to the conductor in the span to change mechanical and/or damping characteristics. Specially shaped conductors or conductor accessories which alter the iced conductor's aerodynamic characteristics, particularly those that increase aerodynamic damping, are also effective. The amplitude of galloping motions can be reduced by the use of higher conductor tensions and evidence suggests higher tensions can also reduce the possibility of occurrence. Galloping and eolian vibration occur in both single and bundled conductors.

Wake-induced oscillation is limited to lines having bundled conductors and results from aerodynamic forces on the downstream conductor of the bundle as it moves in and out of the wake of the upstream conductor. Wake-induced oscillation is controlled by maintaining sufficiently large conductor spacing in the bundle, unequal subspan lengths, and tilting the bundles.

Eolian Vibration. As wind blows across a conductor, vortices are shed from the top and bottom of the conductor. The vortex shedding is accompanied by a varying pressure on the top and bottom of the conductor that encourages cyclic vibration of the conductor perpendicular to the direction of wind flow. The frequency at which this alternating pressure occurs is given by the expression

$$f = 3.26 \times \frac{U}{d} \tag{14-80}$$

where U = wind speed, mi/h
d = conductor diameter, in
f = frequency, Hz

For a 1.0-in-diameter conductor exposed to a 10-mi/h wind, the vortex shedding force oscillates at 32.6 Hz. To develop significant amplitudes, there must be a resonance between this oscillating wind force and the vibrating catenary (conductor). The fundamental frequency of vibration of the suspended conductor is in the range of 0.1 to 1.0 Hz. Therefore, the eolian vibration force will be unlikely to excite a fundamental span mode. This is verified by actual conductor performance where significant amplitudes are usually observed for frequencies in the range of 10 to 100 Hz. Practical wind speeds cause vortex shedding forces of greater than 10 Hz, eliminating frequencies below this level, and frequencies above 100 Hz are not present because of the rapid increase in conductor self-damping for these higher frequencies.

The maximum alternating stress resulting in strand fatigue normally occurs at the conductor clamp. The stress is related to the amplitude of conductor vibration and is the amplitude normally measured by field recording devices. Stress and amplitude of vibration can be related by analytical means such as the Poffenberger-Swart formula.[67] The amplitude of eolian vibration is fixed by the balance of energy input from the wind-induced vortex shedding forces and the energy loss due to conductor, accessory, and structure damping. The addition of dampers to the conductor has been established as an effective means of control.[68] Special conductors such as SDC and SSAC[69] have also been shown effective in reducing the strand stress levels.

Another effective means of limiting vibration fatigue problems is to increase the self-damping of standard conductors by reducing tension. As a practical approximation, stringing conductors to a final unloaded tension of 15% or less at the minimum seasonal average temperature (usually 0 to 30°F) will prevent vibration fatigue problems. Higher tensions are routinely used in areas where the line is parallel to existing lines and the higher tension on the existing line has not resulted in problems. The use of vibration dampers or special antivibration conductors can also allow the use of higher tension levels.

As with single conductors, bundled conductors are subjected to eolian vibration. However, the interaction of conductors in the bundle due to slightly different tensions and increased damping from spacers results in lower vibration levels for bundles than for single conductors in the same wind exposure.

Galloping. Both bundled and single conductors are subject to galloping during or after glaze ice storms. Galloping oscillations occur at frequencies near the fundamental span mode or its second or third harmonic (0.1 to 1.0 Hz) and exhibit maximum amplitudes as large as the conductor sag. While there has been extensive debate concerning the galloping mechanism, and considerable experimental and analytical study, it appears that there presently exist a number of control methods that are effective in reducing the amplitude and incidence of galloping motion. In-span hardware, such as the "detuning pendulum" developed by EPRI,[70] and the wind damper" developed by Richardson,[71] are effective for existing spans where galloping occurs. The T2 conductor,[72] developed by Kaiser, and several other hardware devices are available to control galloping in new lines.

In contrast, control methods such as sleet melting by use of high current levels appear to be almost totally ineffective in stopping galloping and can result in annealing damage to the conductor.

Wake-Induced Oscillation. Bundled conductors are subject to wake-induced oscillations with amplitudes and frequencies typically between that of eolian vibration and galloping. The frequencies of oscillation are normally in the range of 1 to 10 Hz, and the amplitudes are in the range of 10 conductor diameters. The modes in which such vibration occurs are considerably more complex than the modes exhibited during either galloping or the almost invisible eolian vibrations. The source of wind energy for wake-induced oscillation is, as the name suggests, the wake from the windward conductor of the bundle which causes the motion of the downwind conductor.

There are three basic approaches to the control of wake-induced oscillation.[73] Two involve reducing the input of wind energy, and the third involves detuning the mechanical bundle system to prevent resonance. The methods based on reducing wind energy input to the bundle are bundle tilting and bundle sizing. By tilting the bundle to angles of 20° or more, the downwind conductors are moved to the edge of the upwind conductor's wake and the energy input is reduced. By keeping the subconductor spacing to the order of 20 times the conductor diameter, the wind energy input to the windward conductor is reduced by being moved to a wake region of reduced intensity. The third commonly used method to control or eliminate wake-induced oscillations is to stagger the length or simply to shorten the average subspan length. This method does not control those oscillations where the bundle moves as a rigid body and is somewhat dependent on the mechanical characteristics of the spacers.

In comparison to the damage that can result from eolian vibration or galloping, field reports of wake-induced oscillation damage are usually of a minor nature, primarily conductor abrasion from clashing and spacer breakage, neither of which normally results in system outages.

14.1.9 Supporting Structures

Types of Supporting Structure. Numerous types of structure are used for supporting transmission-line conductors, for example, self-supporting steel towers, guyed steel towers, self-supporting aluminum towers, guyed aluminum towers, self-supporting steel poles, flexible and semiflexible steel towers and poles, rope suspension, wood poles, wood H frames, and concrete poles. The type of supporting structure to use depends on such factors as the location of the line, importance of the line, desired life of the line, money available for initial investment, cost of maintenance, and availability of material. Because of the wide conductor spacing required for electrical clearances and insulation, the high tensile stresses used in conductors and ground cables to pull these cables up to a sag which will keep the heights of the structures within reason, the long spans necessary for crossing ravines in mountainous country, and the reliance to be placed on a major trunk line, lines exceeding 345 kV are frequently built of self-supporting steel towers although guyed and rope-suspension structures are increasingly applied. A line built with self-supporting steel towers is very satisfactory in all respects, as it requires less inspection and has a maximum life with minimum maintenance costs. However, high-strength aluminum-alloy towers are available, and their use is on the increase. They have the advantage of better resistance to corrosive atmospheres than steel.[74] The structural configurations and design details are the same as with steel, with the added problem of greater deflections when stresses are applied owing to the lower modulus of elasticity of aluminum. The effect of long-time creep of aluminum is yet to be determined. Self-supporting steel poles are frequently used in congested districts where right-of-way is limited and short spans are necessary. The advent of EHV has brought a great variety of new structural configurations. Details of some of these have been published. *Electrical World,* Nov.15, 1965, pp. 95–118, contains outline drawings of 35 towers and six wood-pole H-frame structures as applied to EHV, as well as a tabulation of specification items of EHV lines in the United States and Canada. The *Transmission Line Reference Book, 345 kV and Above,* 2d ed., 1982, published by EPRI,[3] contains details of a broad spectrum of 345- through 800-kV structures.

Wood poles are used extensively where they are readily available. Medium- and lower-voltage lines can be built economically with such poles fitted with either steel or wood crossarms. *Wood* H *frames* composed of two poles tied together at the top with wood or steel crossarms have been successfully used for the higher-voltage lines up to 345 kV. To take full advantage of the transverse strength, such poles can be braced internally for at least a portion of their height with wood X bracing.

Concrete poles have been used in some parts of the world where timber is scarce and where the ingredients for making concrete are readily obtainable. Another advantage is that they are impervious to insect damage and other forms of decay prevalent with wood structures in tropical or subtropical climates. They are generally cast in units, by using standard forms, and transported to the site, although they may be manufactured where used. Concrete poles should always have sufficient prestressed steel reinforcement to take care of the bending stresses due to wind loads, pulls from cables, and the like, in addition to being designed as columns under vertical loads. In all structures conductor configuration and the effect of various forces which may act upon them must be taken into account.

Conductor Spacing and Clearances

Horizontal Configuration. The minimum spacing of conductors on structures where post-type or V-string insulators are used on medium-length spans will generally depend on the least separation that can be used at midspan without the conductors approaching too closely under adverse wind- or ice-loading conditions.

With suspension insulators a different problem exists, as the swing of the insulator string has to be considered and clearances to the structure determined. This will generally give conductors a spacing at the supports which will be greater than the required midspan separation. One typical rule is to calculate the swing of the insulator string, both with the wind on the bare conductor and the wind on the ice-coated conductor with the corresponding vertical loads acting at the point of conductor suspension, to determine which condition gives the maximum deflection. The vertical loads should be taken on a length of span which is two-thirds the span for the horizontal loads. This will allow a certain amount of leeway in using a standard height structure at a location where the ground is lower than at the two adjacent structures. After the length of the insulator string has been determined electrically and the angle of insulator swing calculated, a normal electrical clearance is established to the structure from the deflected position of the conductor, which, when applied to the three conductors in their relative positions, will determine the necessary horizontal separation of the phases at the supports. This separation should then be checked to see whether it is sufficient for the midspan separation required. Midspan separations that will not be subject to flashover if the conductors begin to swing out of step are usually inherent on high-voltage lines owing to the clearances required at the structures. On very long spans and on the longer spans of low-voltage lines, these spacings may be insufficient. Thomas[75] proposed a horizontal-spacing formula for the determination of safe midspan spacings in windy territory where gusts and strong eddies might cause wires to start swinging at different periods

$$\delta = \frac{CdD}{w} + A + \frac{L}{2} \tag{14-81}$$

in which δ = horizontal spacing in feet; C = an experience factor discussed later; d = percent sag of the condition to be studied; D = overall diameter of the conductor; w = conductor weight, in pounds per foot, used in calculating d; A = arcing distance of the line voltage (1 ft/110 kV); and L = length, in feet, of the swinging portion of the insulator string. Thomas proposed an experience factor of 4 for copper and 3.5 for ACSR.[75] It has since been found that, in areas not subject to frequent violent winds, values of C as low as 1 will provide safe midspan spacings. Thomas was doubtful whether as to the added $L/2$ distance is necessary, since insulators seldom swing out of step. This doubt seems to have been justified.

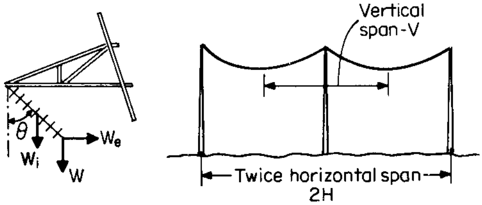


FIGURE 14-29 Determination of suspension insulator swing.

Vertical Configuration. Where the conductors are arranged in vertical configuration, the same electrical clearances will apply for the same voltage as for horizontal configuration, but it may be necessary to increase the vertical separation somewhat to prevent the conductors from coming together or approaching too closely at the center of the span when unequal ice-loading conditions occur or the ice falls off a lower conductor first.

In Fig. 14-29, θ = angle of insulator swing from vertical, H = horizontal span, V = vertical span, w = weight of conductor with or without ice load per lineal foot, and w_i = weight of insulator string including hardware. Then

$$\tan\theta = \frac{Hw_e}{Vw + w_i/2} \tag{14-82}$$

Ground wires, if used, are located above the conductors for lightning protection and in such a position that there is no danger of contact with the conductors at midspan. As ground wires are generally strung with less sag than the conductor cables, ample clearance at midspan is readily obtainable.

These considerations, taken together with the maximum vertical sag to be used and the height required for the conductors above ground level, will determine the height and width of the supporting structure. Extensions can be used where the terrain requires a higher structure than normal.

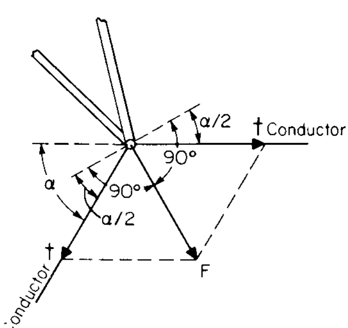


FIGURE 14-30 Determination of transverse forces.

Transverse Forces on Support Structures. Transverse forces acting on towers or poles are due to wind on the conductors and ground cables (and ice coating if in ice districts), wind on the structures, and horizontal components of the tensions in the cables at angle turns in the line (Fig. 14-30).

The stress due to an angle in the line is computed by finding the resultant force produced by the wires in the two adjacent spans. For example, in Fig.14-30, if the change in the direction of the line is the angle a and the stresses t in the adjacent spans are equal to each other, the resultant force

$$F = 2t \sin\frac{a}{2} \tag{14-83}$$

TABLE 14-22 Resultant Force Due to Equal Tensions of 1000 lb in Adjacent Spans

Angle			Angle		
a deg	*a*/2 deg	Resultant *F*, lb	*a* deg	*a*/2 deg	Resultant *F*, lb
10	5	174.4	70	35	1147.2
20	10	347.2	80	40	1285.6
30	15	517.6	90	45	1414.2
40	20	684.0	100	50	1532.0
50	25	845.2	110	55	1638.4
60	30	1000.0	120	60	1732.0

Note: 1 lb = 4.448 N

Table 14-22 gives the resultant force *F* due to a tension *t* of 1000 lb in each conductor of two adjacent spans. The resultant force due to each conductor may be thus computed and the moments about the ground line determined. These moments may be added to those produced by the wind pressure to find the maximum stress.

In applying wind loads to the structure, the appropriate force coefficients, exposure coefficients, and gust response factors should be used.

Longitudinal Forces on Support Structures. Longitudinal forces acting on towers or poles are due mainly to the maximum tension which is assumed to exist in the conductor and ground wires if broken. Ordinarily, especially with suspension insulator strings, these tensions are balanced in the adjacent spans; but if a conductor breaks, a distinct force is produced along the line due to unbalanced tension. If the break occurs on a conductor at the end of a crossarm, there is, in addition to the longitudinal force, a torsional force introduced which must be resisted by the structure. Wind acting in the direction of the line is not ordinarily a factor, as the maximum tension in the conductor is produced when the wind is blowing transverse to the line. As to the reduced stress which occurs in a span from the breaking of a conductor with the suspension insulator string deflecting in the direction of the line, the best practice is to ignore this reduction in tension, as the force due to breaking may cause an impact which more than offsets the reduction in tension. Special release clamps were devised for use on the Plymouth Meeting–Siegfried line of the Philadelphia Electric Company so that, if an insulator string deflected to an angle of 20° in the direction of the line, the clamping mechanism would release the pressure with only the friction in the saddle holding the conductor. This reduced the tension in the conductor considerably; and by assuming a low value for the tension in the conductor due to a break, a more economical structure was obtained.

Vertical Forces on Support Structures. Vertical forces acting on towers or poles are those caused by the weight of that portion of the conductors, plus ice loading if any, which is supported by the structure in question. In addition, there are the weights due to insulators and accessories and the weight of the structure itself. If a structure is located in a valley, there may actually be uplift on it, if the vertical components of the tensions in the conductors exceed the downward loads.

Combined Forces on Support Structures. In determining the maximum forces acting on towers or poles, it is necessary to combine the transverse forces, longitudinal forces (including torsion), and vertical forces so that they act simultaneously. Several different combinations of loading conditions may be desirable, as follows:

1. A condition with all conductors intact and the full transverse and vertical forces acting. These forces should correspond to the appropriate extreme wind and ice loadings and the NESC district loadings. The NESC also specifies overload capacity factors which must be applied to the transverse, vertical, and longitudinal loadings to provide adequate strength of the support structures. These factors depend on the grade of construction and on the type of structure. ASCE Manual 74[57] specifies load factors that can be applied to the extreme wind and ice loading cases to increase the reliability for important or long transmission lines.

2. A condition with all conductors intact, except the number it is desired to assume broken, with the transverse and vertical forces computed for each particular conductor, according to whether it is assumed broken. The longitudinal forces due to broken conductors must be combined with the transverse and vertical forces at all points of support where the conductors are assumed broken. It is customary, when more than one conductor is assumed broken, to consider all breaks in the same span and at the supports which will produce the maximum overturning moment, the maximum torque, or a combination of both.

3. A condition in some localities where extraheavy vertical loads, caused by an unusually large formation of ice on the conductors, may occur. These loads are combined with the weight of the structure.

4. A condition with vertical loads acting upward at the conductor supports.

NOTE: It is not customary to combine transverse and longitudinal loads with the loads specified under items 3 and 4.

Other factors may enter into the determination of the maximum forces acting on supporting structures in special cases, such as the horizontal and vertical components of tensions in guys and the addition of pole-top transformers, switches, and working platforms.[76]

The proper number of conductors to assume broken is a debatable question and depends upon what margin of safety is desired and the amount of money it is desired to invest for this security. Generally speaking, the minimum number of conductors to assume broken for tangent suspension single-circuit towers should be either one ground wire or any one conductor, and for double-circuit towers either one ground wire and one conductor or any two conductors on the same side of the tower and in the same span, by using the different cable supports for application of the forces to determine the maximum stress in each member of the tower. Anchor or dead-end towers should be able to withstand all or any number of conductors and ground wires broken. Generally, the condition of broken conductors and ground wires on one side of the tower will produce greater stresses in the web members than if all the conductors and ground wires are considered broken, owing to the unbalanced torsional forces existing when only the conductors and ground wires on one side of the tower are broken.

Types of Metal Structure. Structures may support single, double, or multiple circuits. The first two types are generally used for transmission lines except in congested areas where right-of-way is very expensive and it is desired to transmit large blocks of power over one line. In such a case three or more circuits may be supported by the structures.

Self-Supporting or Rigid Structures. On both single- and double-circuit tower lines of any considerable length, at least three kinds of towers are required for economic reasons:

1. A tangent suspension tower which can be used for normal spans where no angles in the line occur (Figs. 14-31 and 14-32).

2. An angle suspension tower which can be used for normal spans with a small-angle turn in the line or with longer spans on tangents.

3. An angle tower which can be used for normal spans with a large-angle turn in the line, with extra-long spans on tangent, or as a full dead-end tower for anchoring. Insulators may be either suspended or in the strain position.

Very often it is desirable to introduce a fourth kind of tower with insulators always in the strain position to take care of exceptionally large-angle turns in the line; in extremely long spans on tangent; and also, where required, as a full dead-end tower. When this type of tower is provided, the tower listed in item 3 may be of lighter construction and not used for dead-end purposes.

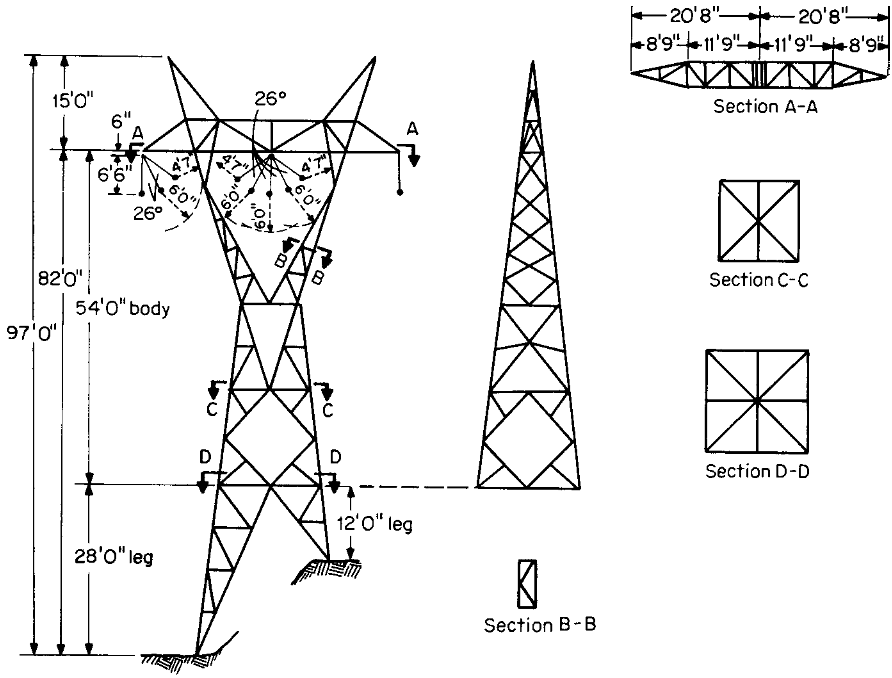


FIGURE 14-31 Tennessee Valley Authority 161-kV single-circuit tangent suspension corset-type tower. (*Designed by Blaw-Knox Co.*)

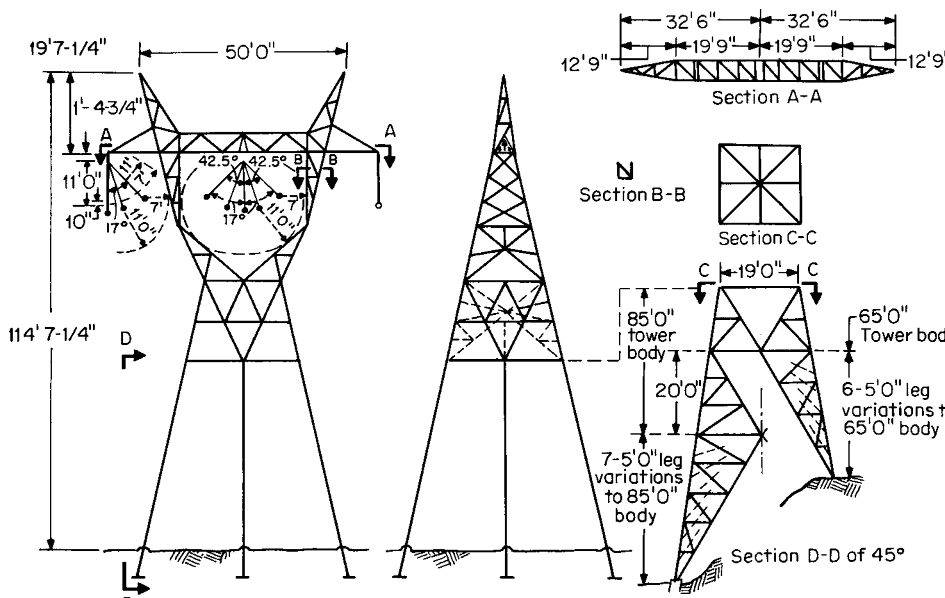


FIGURE 14-32 City of Los Angeles 287-kV tangent suspension rotated-type tower. (*Designed by American Bridge Co.*)

Double-circuit towers with the vertical configuration of conductors, as used on different lines, are very much alike in appearance, generally being square in cross section. It is customary to locate the middle conductors outside the upper and lower conductors.

With single-circuit towers and the conductors arranged in horizontal configuration, a different problem arises which has resulted in the design of special patented structures for the higher-voltage lines with wide conductor spacing. The shape of these towers has been developed with a view to minimizing the weight of steel required in the superstructure and also reducing the size of footings by minimizing the effect of torsion. The more common types are the Blaw-Knox tower (Fig. 14-31), or corseted type, as originally used on the Plymouth Meeting–Conowingo line of the Philadelphia Electric Company; and the American Bridge Company's rotated tower (Fig. 14-32), used on the first Hoover Dam–Los Angeles line (this line has since been uprated to 500 kV) and also by the Bonneville system and on lines of the Tennessee Valley Authority. Either of these types serves the purpose for which it was intended. The theory behind the rotated tower is that the greatest overturning moment is caused by a combination of the transverse forces and longitudinal forces, due to broken conductors, acting simultaneously, which produces a resultant force acting at an angle of approximately 45° with the direction of the line. In this case the whole four tower legs are resisting the overturning moment, thereby reducing foundation loads and consequently costs. Under normal conditions of loading, with only the transverse forces acting, the legs on the diagonal separation will take care of the overturning moment. Obviously the greatest advantage of the rotated type over the nonrotated type is on tangent towers and towers used for small-angle turns in the line when the transverse and longitudinal forces are approximately equal.

Figure 14-33*a* shows a TVA 500-kV conventional-design tangent self-supporting tower for a bundle-conductor line having three 971,600-cmil ACSR conductors per phase. The overhead ground-wire clamps are suspended and insulated from the tower by means of distribution-type guy strain insulators. The overhead ground wires are composed of seven strands of No. 9 Alumoweld and are used for carrier-current communication channels.[77] Each ground-wire insulator is provided with a spill-over gap to protect it during lightning discharges.

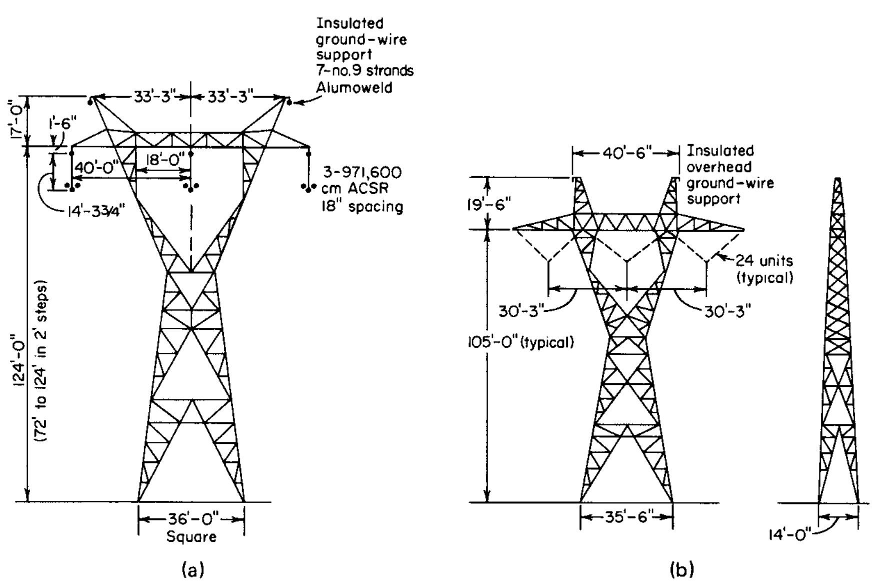


FIGURE 14-33 (*a*) A 500-kV tangent self-supporting tower (Tennessee Valley Authority); (*b*) 500 kV semiflexible steel tangent tower (Arkansas Power & Light Co.).

It is interesting to compare Fig. 14-33*a* and *b*. Both show 500-kV towers, but Fig. 14-33*b* is designed for a narrower right-of-way. The wind side swing of the conductors in a span is half the sag at 30° side swing, and this is common to both towers. Therefore, the saving in right-of-way for Fig. 14-33*b* is 40 ft plus 7 ft, 7 in, less 30 ft, 3 in, or 17 ft, 4 in on each side, or a total of approximately 35 ft.

Figure 14-34 shows a light-suspension 500-kV single-circuit tower typically used by the Bonneville Power Administration (BPA), supporting three 1,192,500-cmil ACSR Bunting conductors per phase and two 7-strand No. 8 Alumoweld overhead ground wires. BPA uses continuous overhead ground wires throughout its entire 500-kV network except on single-circuit lines west of the Cascade Mountains. In the latter case, overhead ground wires extend 1 mi out from the substations. Typically, BPA 500-kV lines are designed to withstand 100-mi/h winds and solid ice coatings up to 1½ in.

A steel suspension self-supporting tower used by Hydro-Quebec for 735-kV Manicouagan lines is shown in Fig. 14-35. Line conductors consist of a four-conductor

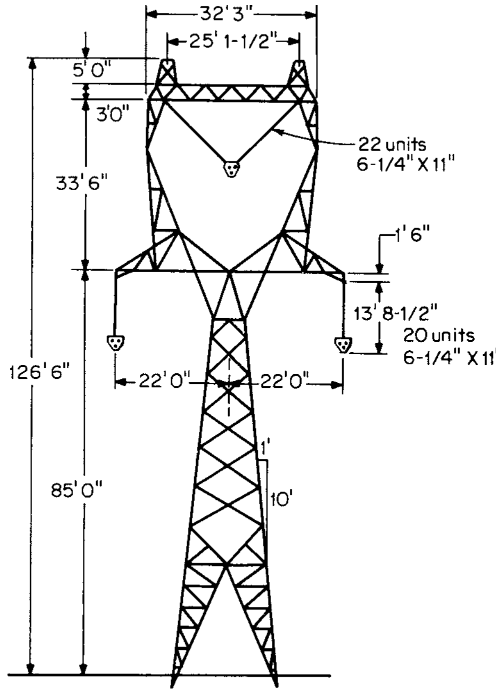


FIGURE 14-34 Suspension-type 500-kV tower. (*Bonneville Power Administration*).

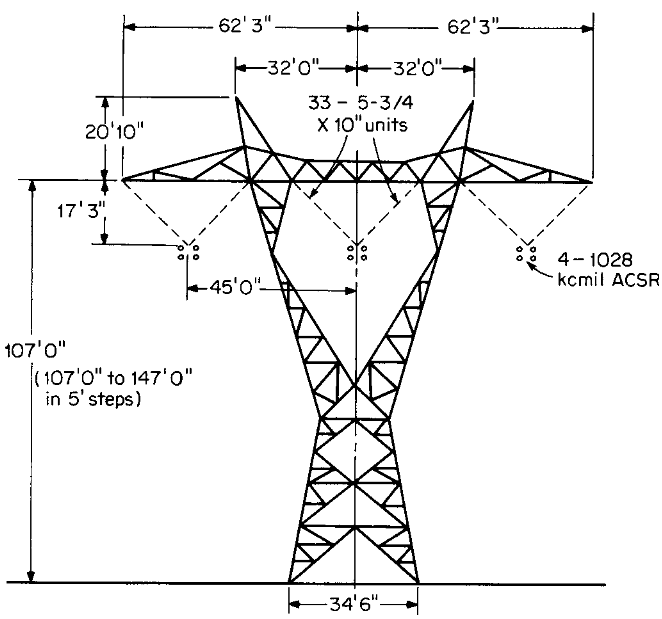


FIGURE 14-35 Steel suspension self-supporting tower for 735-kV Manicouagan lines. (*Hydro-Quebec*.)

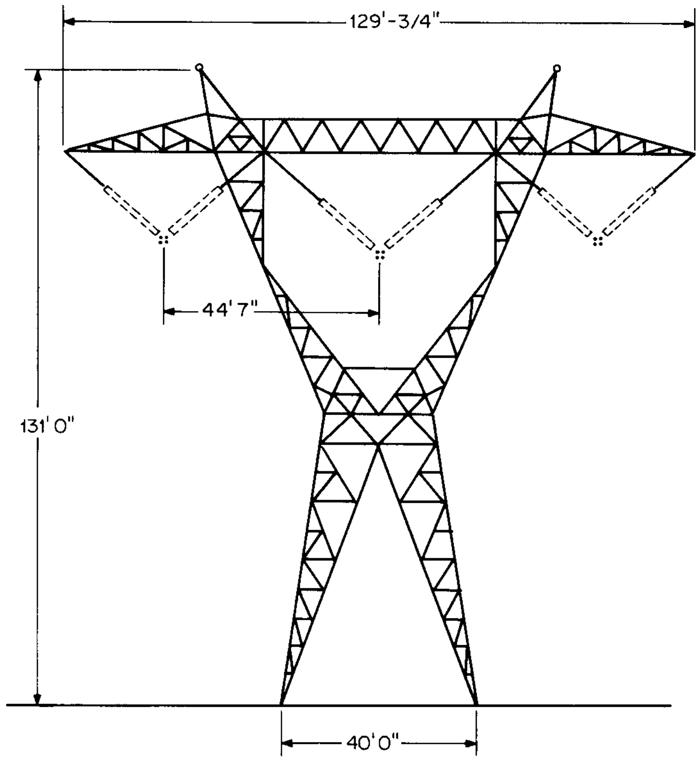


FIGURE 14-36 Steel 765-kV self-supporting suspension tower. (*American Electric Power Co.*)

bundle per phase, where each conductor is a 1,028,000-cmil ACSR insulated with 33 insulator units ($5^3/_4 \times 10$ in) per phase. This type of tower was used on the first stages of the Manicouagan project since September 1965, and in subsequent stages of the same project.

A unique structure is the 765-kV self-supporting steel tower used by American Electric Power Company (Fig. 14-36). This tower, weighing from 44,000 to 66,500 lb, including grillage foundation, was designed by American Bridge Company for erection in parts, if desired, by a Skycrane helicopter. Like AEP's 765-kV V tower shown in Fig. 14-33, there are 30 insulator disks ($5^3/_4 \times 10$ in) per leg of V strings in the outside phases and 32 insulator units in the middle phase. Also, like the tower shown in Fig. 14-33, this tower is designed to meet the same special AEP loading criteria already described. Two overhead ground wires provide a 15° shielding angle to the outside four-conductor bundles.

Semiflexible Structures. Such structures have been used to some extent for the voltages under EHV. This type of tower has a narrow base in the direction of the line. The ground wires are strung tightly to take up unbalanced loads due to broken conductors and form part of the structural system. In case a conductor breaks, the unbalanced load will be taken up by the ground wires and transmitted by them to the next anchor tower.

With the advent of EHV and bundle conductors, semiflexible self-supporting towers are receiving more consideration, and some are being used. With the heavy bundle conductors, the breaking of one conductor is not serious, and the breaking of all conductors of a phase is practically nonexistent. Possible causes are airplanes and tornadoes, which no practical tower could withstand.

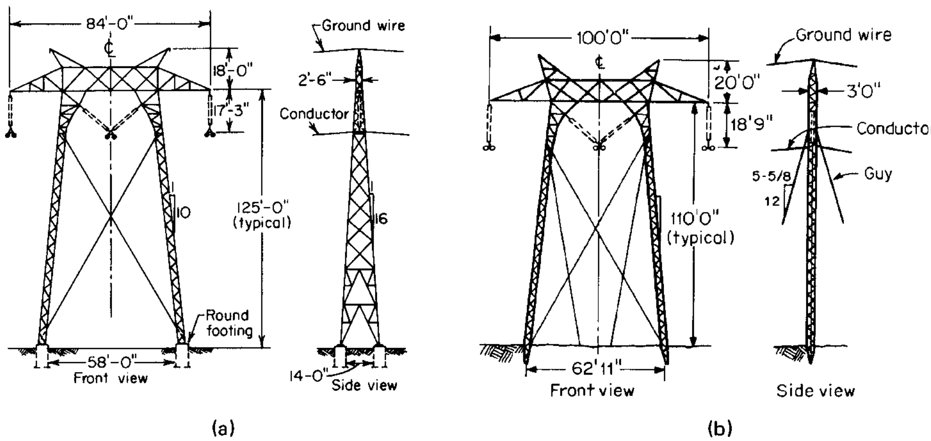


FIGURE 14-37 (*a*) A 500-kV steel tower used in valley areas (*Pacific Gas and Electric Co.*) (*b*) A 500-kV steel tower used in mountainous areas (*Pacific Gas & Electric Co.*)

Figure 14-33*b* shows such a tower as used on the 500-kV system of the Arkansas Power and Light Company. The overhead ground wires are insulated from the towers as they are in Fig. 14-33*a* for communication purposes. Figure 14-37*a* shows a steel-saving semiflexible tower used by the Pacific Gas and Electric Company. Note the X guying used between tower legs to obtain the required lateral strength.

Guyed Towers. Such towers overcome the weakness of semiflexible towers in line with the line. They can be used for single-conductor lines or for any other service. Figure 14-37*b* shows a guyed steel tower used by the Pacific Gas and Electric Company in mountainous country. A feature of the tower is that the legs do not have to be of equal length. This tower has the same internal X guying as the tower of Fig. 14-37*a*. The self-supporting feature of the tower of Fig. 14-37*a* is replaced by four guys in the direction of the line and with an increase in strength.

Figure 14-38 shows a Kaiser aluminum guyed-V 345-kV tower used by the American Electric Power Company. Weighing from 3350 to 5400 lb (1510 to 2450 kg), this tower was erected by using a helicopter to “tilt up” the assembled tower by pivoting about a special hinge at the center foundation. This allows the use of a helicopter with a lifting capacity smaller than the weight of the tower since most of the tower weight is supported by the foundation while it is being tilted up. There are 15 insulator units ($5\frac{3}{4} \times 10$ in) per leg of the V strings on this tower.

Figure 14-39 shows a Kaiser aluminum guyed-V 765-kV tower also used by American Electric Power Company. There are 30 insulator units ($5\frac{3}{4} \times 10$ in) per leg of the V strings in the outside phases and 32 insulator units in the middle phase.

Each of these towers has been designed to withstand special AEP loading criteria which include 100-mi/h winds with no ice, 50-mi/h winds with 1 in of ice, and $1\frac{1}{4}$ in of ice with no wind, in addition to the NESC loading requirements.

Tubular Steel Poles. These poles are being used on city streets and in congested areas where a wide right-of-way cannot be gained. They have been used for voltages up to and including 345 kV. Vertical configuration of conductors is used for all high-voltage lines. Insulators may be side post or suspension[78] on cantilever arms or a combination[79] of the two. Figure 14-40 shows a 230-kV pole used on a line of the Arizona Public Service Company in Phoenix. These poles are of tubular steel in three sections with telescoping joints. The poles are tapered, with a diameter of 24 in at the base and 10.8 in at the top. The mast arms are 8 ft long, of tubular steel, with brackets bolted to the poles with two $\frac{3}{4}$-in through bolts. The poles are spaced approximately 300 ft apart. Insulator side swing is reduced by a 200-lb combined hold-down weight and corona shield. The poles present a pleasing appearance and have elicited no objections even with a line installed on each side of a 60-ft street.

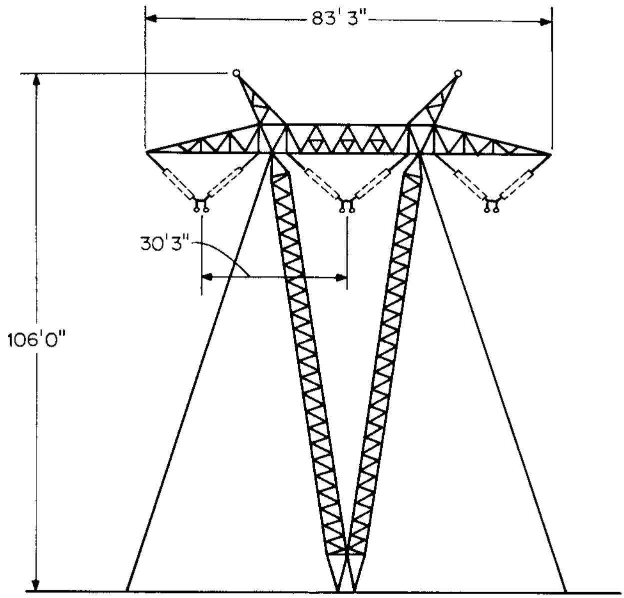


FIGURE 14-38 A 345-kV guyed-V aluminum suspension tower. (*American Electric Power Co.*)

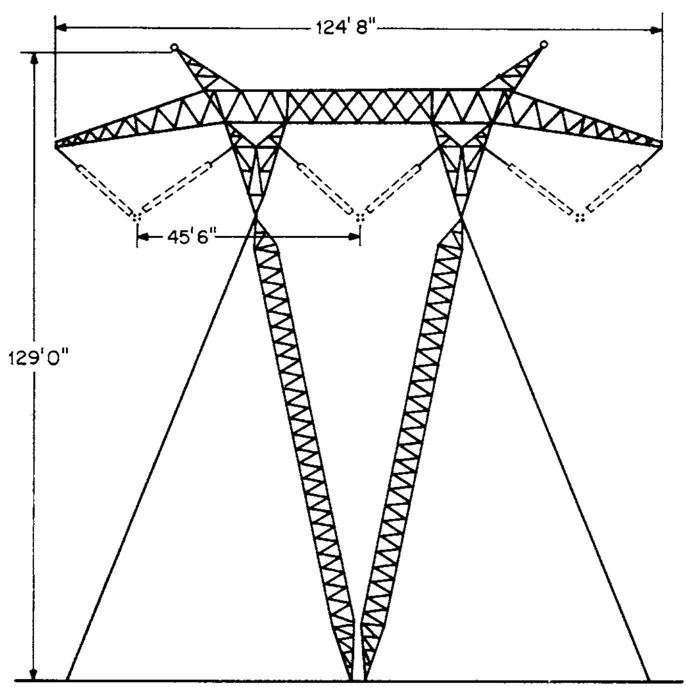


FIGURE 14-39 A 765-kV guyed-V aluminum tower. (*American Electric Power Co.*)

The poles, side arms, and accessories were furnished by the Union Metal Manufacturing Company of Canton, Ohio.

The New Orleans Public Service Company 230-kV line[79] is of similar construction but is designed for hurricane-force winds. The poles are 12-sided, elliptical, high-strength steel, with the short diameter, which is in line with the line, 75% of the long diameter. The insulators are a combination of 12 suspension insulators and a swivel-ended strut (side-post) insulator equal to 12 suspension insulators, to prevent side swing of the suspension insulators. Some poles have side-post 230-kV insulators only. The poles have no base for bolting to a foundation but do have baseplates and are set in concrete in holes 25 ft deep. The holes are made by driving 32-in-diameter steel casings to a depth somewhat deeper than 25 ft and cleaning them out.

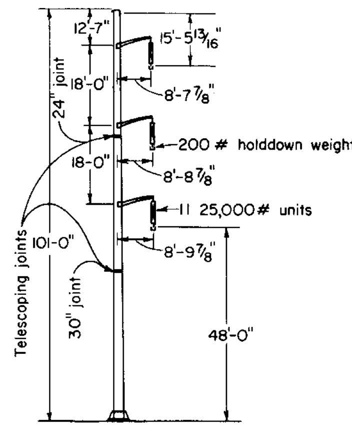


FIGURE 14-40 A 230-kV steel pole. (*Arizona Public Service Co.*)

Special Structures. Transposition towers require special structures where it is not expedient to make the transpositions on a standard tower by the use of special crossarms. Long spans over rivers and bays and crossings over important highways and trunkline railroads frequently require towers which either are much higher than normal or must have a larger factor of safety against collapse. Anchor towers near substations, towers for mounting switches, and towers for turning 90° angles also may come under this classification. Such special structures are designed to suit local requirements and are subject to regulations of the U.S. Army Engineer Corps in the case of navigable-river crossings, to state public-utility commissions or other bodies for highway crossings, and to the particular regulations of railroads which are concerned.

Stresses in Structures—Design. Stresses in towers can be computed analytically by the historic graphic methods or by use of one of the available computer programs for tower analysis and design. Most of the design procedures assume the foundations are rigid. In actual practice, with towers set in the ground, an uneven settlement of the foundations may produce excess stresses which must be considered and taken care of in the overload factor provided against failure. Some flexible lattice-type towers and most pole-type structures will undergo sufficiently large horizontal displacements such that the nonlinear stresses due to the vertical loads acting in conjunction with these displacements (the so-called P-Δ effect) must be considered in the design.

Tension members may be designed for their full net area of section with bolt holes deducted, and compression members may be designed in accordance with the strength formulas given in ASCE Manual 52.[80]

For short panels in narrow-faced towers, it is economical to use a tension and compression system of web bracing without horizontal struts, as the stresses in the corner posts will be reduced considerably, with torsional stresses eliminated entirely. The bracing should make an angle of 30° to 45° with the horizontal, with all web members in any one panel of the same size to divide the loads equally. Where a tension system of diagonal bracing is used with horizontal struts taking the compression, the angle of the bracing may be increased to 50° or more from the horizontal. Long, unsupported lengths of main members can be broken up by the use of redundant members. The lower panels of towers should be so designed that variable-length leg extensions, which are interchangeable, can be employed to take care of sloping ground. Square extensions of any desired length may be used where towers higher than standard are necessary, with variable-length leg extensions fitted to the bottom of the square extensions if required.

Structure Tests. When an important transmission line is built using structures of new design, at least one type of structure (generally the tangent suspension tower or the type which is to be used most frequently) should be tested, first with the working loads as specified for the design of the tower applied, and finally with the ultimate loads which the tower is expected to withstand, applied. Steel poles should always be tested on rigid foundations with the transverse, longitudinal, and vertical loads applied simultaneously to introduce all direct bending and torsional stresses. Crossarms should be tested for the additional torsional stresses introduced, where pin insulators are used, and combined with the longitudinal loads and the heaviest vertical loads specified. Equivalent concentrated loads may be used in some cases to avoid applying a multiplicity of small loads at different points, which would cause delay in shifting loads, but care should be taken to see that all combinations of loads or individual loads which will produce the maximum stress in each member are applied. After the structure has successfully withstood all the specified loads, a destruction test is desirable to determine the overload factor. This can generally be made with the test loads which cause the maximum stresses in the greatest number of members on a tower in place, by increasing the transverse loads indefinitely until failure occurs. After a test is completed, members of a tower should be examined for elongation of bolt holes, straightness, etc.

Towers should be tested with the protective coating which is to be used in service on the steel, and the foundations should be the same as those for which the towers are designed. If it is impossible to test towers on earth foundations, they may be tested on rigid foundations, but a test on rigid foundations will undoubtedly show a greater overload factor than may be expected in service.

Erection of Towers. Towers may be assembled on the ground and then lifted into place by means of self-propelled derrick cranes or latticed-steel gin poles.[81] Very large towers are usually assembled in sections, and the sections are lifted into place by means of cranes or gin poles. Towers in inaccessible locations may be transported and assembled by the use of helicopters. If necessary, the towers can be erected from the ground up by the use of gin poles moved from corner to corner and with the erection crew climbing up on the partly completed structure. This method may be used for small jobs which do not warrant the use of heavy cranes or for very tall towers beyond the reach of cranes, such as those for river crossings.

Insulator and Ground-Wire Attachments. (Fig. 14-41). The method of attaching suspension insulator strings varies. One form of attachment is the U bolt fastened on the underside of the crossarm to which the insulator hardware is attached. This device will give flexibility both longitudinally and transversely. Another attachment is in the form of a bent plate or angle fastened to the underside of the crossarm with a fairly large hole to receive a hook or shackle at the top of the insulator string.

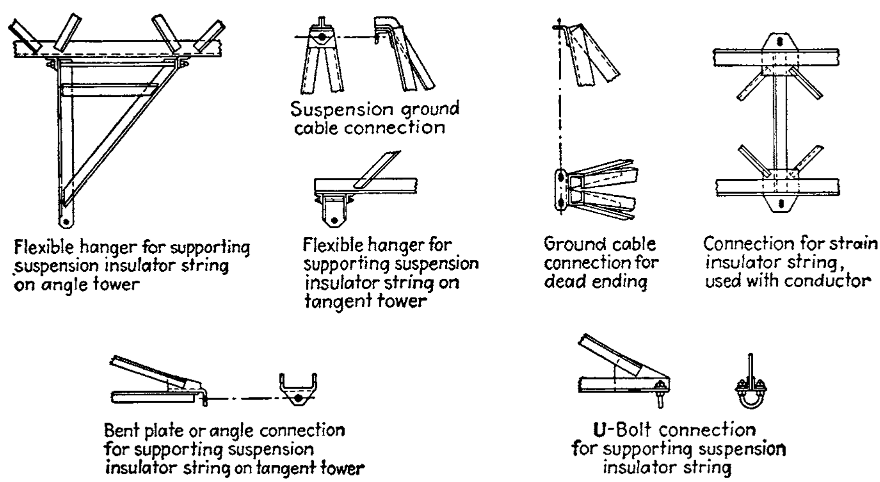


FIGURE 14-41 Insulator and ground-wire attachments.

In order to keep the conductor spacing to a minimum on high-voltage lines and take advantage of clearances to steelwork, the point of attachment of the insulator string may be dropped several inches below the crossarm, in which case flexible hangers are required, that is, hangers which are hinged at the crossarm and free to pivot in the direction of the line. These hangers may be made of plates, shapes, or bent round rods, with suitable connection at the bottom for receiving the insulator string. With suspension-type insulators in the strain position, horizontal pull-off plates are required.

To minimize failure of ground wires due to vibration, they are preferably suspended, and the attachment at the tower consists of an angle or bent plate with a hole to receive the suspension clamp. Patented rigid ground-wire clamps may be obtained if desired, which can be bolted directly to the steel structure. These are generally of the V-groove type.

Ladders and Step Bolts. Ordinary steel towers are provided with step bolts on one corner post for climbing the tower. Special high structures, such as river-crossing towers, should have ladders extending up to the level of the top crossarm. Such ladders should be provided with guard cages supported on the sides of the ladder.

Seismic Effects. This field is complex,[82] involving the various disciplines of the seismologist; the special dynamicist, who understands both structures and equipment and their differences; the vibration test engineer, and the designer who is experienced in the practical aspects of the lines and equipment and their function. There are many subtleties which are rarely appreciated by any one person and, consequently, many authorities must be consulted.

As contrasted with substation equipment, transmission-line structures built to withstand aboveground weather also will generally survive moderate earthquake tremors without noticeable distress. As studies of seismic effects are becoming more sophisticated, investigators have established seismic design criteria in the United States by dividing the country into four zones of seismic probability, the most severe being on the Pacific coast.[83] In such areas periodic review of existing foundations from a seismic standpoint is recommended. Experience in the San Fernando, California, earthquake of February 9, 1971, showed that over the years the foundation strength of a number of towers had been sufficiently reduced by erosion or adjacent excavation that their slopes failed during the earthquake.

One study[84] analyzes the effect of earthquake ground motion on both wood and steel transmission structures. It concludes that, except where damage to foundations and anchors occurs because actual earth fissures or slippage develop, seismic disturbances produce no overstress in transmission structures designed in conformity with the *National Electrical Safety Code.*

Protective Coatings for Steel Structures

Galvanizing. For important transmission lines where long life is desired, it is almost universal practice to galvanize fabricated-steel towers and poles which are field-bolted after the other shop work has been completed, because under such conditions galvanizing is more economical than painting. The method of "hot-dip" galvanizing is also used for bolts and nuts, with the threads rerun for nuts after galvanizing. This has practically superseded the sherardizing, or "dry galvanizing," of bolts and nuts which was in use for a number of years. All galvanizing should be in accordance with the *Standard Specifications for Zinc (Hot Galvanized) Coatings on Structural Steel Shapes, Plates and Bars and Their Products*, as given by ASTM Designation A 123, which calls for an average coating of 2 oz of zinc per square foot of surface.

To test the uniformity in the thickness of the galvanized coating, the Preece test is used. This is described in the *Standard Methods of Determining Weight and Uniformity of Coating* in *Zinc-Coated (Galvanized) Iron or Steel Articles*, as given by ASTM Designation A90.

Structures located near industrial plants, if subjected to sulfuric acid and fumes, should not be galvanized.

Painting. Painting is sometimes resorted to for fabricated-steel towers and poles, and generally shop-riveted, welded, or special steel poles are painted. Towers located near industrial plants in a smoky atmosphere should be painted. The base coat should be a mixture of red lead and raw linseed oil in something like the proportion of 33 lb of dry red lead to a gallon of oil with a little dryer added. The outer coats may be of any good all-weather paint. To keep structures in good condition, painting is necessary every 2 or 3 years. With structures having a larger number of small members, this may make the cost of maintenance very high. On structures having few pieces and large, flat surfaces, painting may be economical.

Where steel structures are buried in the ground, a special problem sometimes arises at the ground level where moisture occurs. At this point, galvanizing may deteriorate after a short interval of time, especially if the soil has a sulfur or acid content. A paint made from asphaltum compounds will often prove useful for protection at these points.

Newly galvanized steel should not be painted until it has a chance to weather for a period of 6 months, and the galvanized surfaces should then be cleaned. Aluminum paint is ordinarily used to paint galvanized towers when the galvanizing has deteriorated.

Weathering, or self-painting, steels have been developed by the major steel companies. These are steels which are not treated with any kind of protective coating, but the chemical composition of which is such that they may be said to "paint" themselves. They are installed completely uncoated but thoroughly cleaned of all mill scale and foreign matter. In a few years, a dense dark-brown oxide with a purplish cast forms which becomes a permanent protective coating to all surfaces exposed to the weather. A slight loss of thickness occurs, which eventually stops, as the corrosion rate is nonprogressive.

Wood Poles. Wood poles are considerably cheaper than steel for many types of construction. The lower cost is due, in part, to the more conservative basis of design normally adopted for steel. Generally, steel structures are designed to support safely one or more broken conductors, whereas wood structures are often not so designed. It is logical that the reasons for choosing the more expensive steel construction should require conservative design throughout and that conditions justifying the cheaper and shorter-lived wood structures would warrant accepting some of the more theoretical hazards.

For voltages of 69 kV and lower, wood is quite generally used.

Wood-pole construction for many years has been used for all voltages up to and including 345 kV. H frames with various modifications have been designed, the most popular using the main crossarm as the bottom member of a truss.

Butt-treated cedar and full-treated pine are used almost exclusively in transmission-line construction; the use of untreated poles has been practically abandoned as uneconomical since the supply of chestnut and northern cedar poles has been exhausted. Treated fir has also been supplied in some quantity from the Northwest.

Cedar poles resist decay, but satisfactory life is not secured unless the butt is treated. The pole is usually treated from the butt to about 2 ft above the ground line. The balance of the pole is not treated. Pine and fir require complete treatment of practically all the sapwood. This treatment is applied under pressure.

No universally effective protection has been devised against woodpecker damage. Some localities are often subject to serious epidemics of woodpecker trouble.

Preservative Treatment. Pole decay is due to a fungus which requires air, moisture, warmth, and food for its subsistence; the wood of the pole constitutes its food. The conditions most favorable to the growth of the fungus are found at the ground line. The preservative has toxic or antiseptic properties which make the wood either poisonous or unfit food for the fungus.

Preservatives and preserving methods conforming to the standards of the American Wood Preservers Association (AWPA)[85] should be used in the treatment of poles. There are many wood preservatives, including those using poisonous salts such as copper, mercury, zinc, and arsenic compounds. However, only two are included in AWPA recommendations for poles, Standard C-4-74-C:

1. Coal-tar creosote, AWPA Standard P1-65
2. A 5% solution of pentachlorophenol in a petroleum distillate, AWPA Standard P8 (commonly called "penta")

By AWPA Standard M1-70, pentachlorophenol is not recommended for use in coastal waters. Coastal waters are defined as salty waters. One other preservative is increasing in popularity. This is

AWPA Standard P11-70, a creosote-pentachlorophenol mixture in which pentachlorophenol is not less than 2% of the mixture. All of these preservatives are applied by the following methods:

1. *The open-tank method*, applied to cedar poles, consists in boiling the butts of the poles in a tank of creosote oil, after which the oil is allowed to cool or the poles are transferred to a cold tank of oil. The duration of the hot and cold treatment, usually 8 h or more, depends on several factors, the most important of which is the degree of seasoning. The treatment is based on the fact that the wood cells expand with heat and on cooling draw the creosote into the wood under atmospheric pressure. The sapwood of unseasoned poles has annular rings of a nearly impervious fiber which prevent penetration of the oil. In seasoning, this fiber dries and breaks open. To ensure penetration of the greater part of the sapwood, which is usually less than 1 in in depth, an incision process has been developed and is almost universally used. Narrow cuts, parallel with the wood fibers, are made to a depth of about 1/2 in at frequent intervals around the circumference of the pole for a distance above and below the ground line. Complete penetration is obtained to a depth somewhat greater than the depth of the incisions even on unseasoned poles.
2. *Pressure treatment* is applied to pine and fir. The poles, on a truck, are run into a steel cylinder and subjected to a steam treatment for a period of several hours at a temperature which will not damage the wood cells, usually specified at not more than 259°F (126°C). The pressure is then removed and a vacuum applied. The steam treatment opens up the wood cells and allows the preservative to penetrate. The length of time required for the steam and vacuum treatment depends on the condition of the wood, the amount of oil that is to be injected, and the depth of penetration desired. From this point in the process, one of two methods may be followed. The *full-cell*, or *Bethel, process* allows all the preservative injected to remain in the wood. This process is generally used for piling and underwater work when it is desired to exclude water from the wood and to resist the attack of marine borers. The *empty-cell process* draws off excess oil and secures protection from decay by the coating of oil left on the walls of the wood cells. The empty-cell process is adequate and preferable for usual structures and is used almost exclusively for poles and arms.

The empty-cell treatment is obtained by either the Rueping or the Lowry process. The Rueping process seems to be in more general use, although the Lowry process is equally successful.

In the *Rueping process*, following the steam treatment, an air pressure is applied. While still under pressure, hot oil is forced into the cylinder. The oil is held under this pressure and maintained at a temperature of about 200°F (93.5°C) by steam coils within the cylinder, for a period of several hours. Upon removing the oil and reducing the pressure, the compressed air within the wood cells forces out the surplus oil. The amount of oil retained depends on the pressures applied and the time of treatment, although it is possible to remove only a part of the oil that has been injected.

The *Lowry process* is similar to the Rueping process except that no compressed air is used. After the preservative has been forced into the wood under pressure, a high vacuum is quickly created, causing a sudden expansion of the air within the wood cells and thus driving out surplus preservative (see also Sec. 4).

Strength Calculations. As used in a line, the pole is a cantelever beam, fixed in the earth at the butt and supporting the transverse wind load from the conductors of a length equal to half the sum of adjacent spans. Computation of the safe load that may be carried is a matter of simple mechanics outlined in Fig. 14-42. Some slight approximations have been introduced for simplicity.

The fact that if the pole were a part of a perfect cone, the maximum fiber stress might occur at a point above the ground line is of more theoretical interest than practical use. The difference between the load carried at the critical section and at the ground line is less than may readily be caused by irregularities in the pole.

Poles are almost universally classified according to the ANSI dimensions (see Sec. 4), which have been arranged so that the nominal ultimate strength is the same for all lengths and species of the same class. Poles are classified as Class 1, 2, 3, etc., and H1, H2, H3, etc., and the minimum circumference 6 ft from the butt is specified for each class and each species to give the desired nominal strength. The nominal ultimate was computed from conservative average ultimate fiber stresses from a very large number of tests.

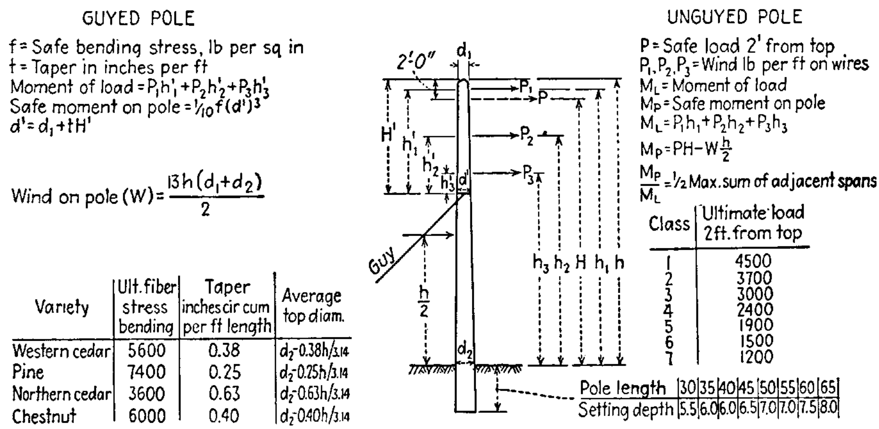


FIGURE 14-42 General pole-strength calculations.

The top diameters are specified but are given only as a minimum and are the same for all species. Actually, the taper of various kinds of timber, although fairly uniform, is quite different for different species, and the average top diameter of ANSI-class poles will be considerably larger than this minimum.

1. *Pole tests.* These give very erratic results, and tests on a few poles should never be given great consideration. Designs should if possible be based on accepted average unit fiber stresses rather than test results unless a considerable number of duplicate tests can be made and averaged.
2. *Factor of safety.* It has been found from experience with heavy transmission-line construction that a safety factor of 2 on the accepted average ultimate is conservative. On light construction, this is sometimes slightly reduced, but a material reduction is usually not justified in view of the deterioration of wood with age. On sustained loads, such as heavy angles, a liberal additional factor of safety is desirable to prevent the pole's warping and giving the appearance of being overloaded. When possible, guys should be attached close to the load to eliminate heavy continued bending.

Setting Depth. The strength of the pole foundation is difficult to reduce to figures and is not of such primary importance to the safety of the structure as in the design of a tower. Failure of the foundation, in the sense that failure is used in the design of steel towers, that is, a considerable movement of the pole in the ground, is of little consequence except for the inconvenience and expense of straightening up the line and retamping the poles. The setting depth for poles of various lengths has been pretty well established by general practice and is almost universally used (see Fig. 14-42). These depths seem somewhat illogical in that no account is taken of the strength of the pole or of the quality of the soil; however, this appears more reasonable when it is considered that the desired result is not to obtain a rigid foundation but to prevent the pole from "kicking" out of the ground.

Wood Crossarms. These are generally manufactured of creosoted yellow pine or untreated Douglas fir. Untreated pine arms of the timber commercially available are not satisfactory. Untreated fir arms are widely used and are apparently giving a life comparable with that of the poles. Arms should be of the highest-quality timber. The smaller arms, up to 5 × 6 in and 10 or 12 ft in length, can generally be supplied on standard crossarm specifications, although structural specifications give very satisfactory arms. Heavy arms for H frames, that is, 6- × 8- and 6- × 10-in timbers and 3- × 8- and

3- × 10-in planks, 20 to 35 ft long, are best purchased as high-grade structural timbers. Structural timbers are furnished under the rigid specifications and inspection of the large timber manufacturers' associations.

Plank Arms. The eccentric connection of large arms to the pole, especially when carrying heavy conductors, is not desirable, and a number of designs make use of two planks, one on each side of the pole attached together at the ends. Generally two 3- × 8-in planks are used in place of a 6- × 8-in timber arm. The plank-arm construction has several advantages in addition to the better connection to the pole, although the end hardware is somewhat complicated, and in many designs the strength of the crossarm against longitudinal loads is somewhat reduced.

Wood versus Steel Arms. Wood crossarms are lower in cost than steel arms of the same strength and, aside from the shorter life, the possibility of being shattered by lightning, and the risk of burning due to leakage current at 345 kV and above, are satisfactory. On wood pole construction, the advantages of steel crossarms—resistance to lightning damage and longer life—are not usually sufficient to offset the insulation strength of wood crossarms. In any event, lightning damage to arms is usually not a major operating problem; and on lines thoroughly shielded with overhead ground wires, crossarm and pole damage is practically eliminated. At higher voltages, crossarms are sometimes equipped with bonding wires to prevent leakage current from burning.

Design of Arms. The design must provide for carrying the vertical load with an ample margin of safety, but often neither the arm nor the connecting hardware is well suited for carrying the full load of a broken conductor as is required of steel towers. Crossarms on single-pole construction have practically no resistance to longitudinal loads. If a heavy conductor breaks, the arm will swing around to very nearly a longitudinal position, restrained only by the attachment of the unbroken conductors. This would be likely to result in badly bent hardware and probably a split and disfigured arm but little serious damage; the major damage would be the broken conductor and not the effects of the break. H-frame construction (see Fig. 14-46*c*) is better adapted to such loads, but the effect of a break is much the same, in that the deflection of the poles and movement of the poles in the ground relieve the greater part of the load and usually result in only some minor damage to the arms and hardware, which is easily repaired.

Double-arm construction can be considered very little, if any, stronger than twice the sufficient bolts and keys to develop the shear. The shear is several times the applied load and makes a very heavy joint necessary.

The common sizes used, 5 × 6 in for lighter single-pole construction and 6 × 8 in for H-frame, allow ample vertical strength for ordinary spans. Conservative unit stresses should be used in vertical load on the arms.

The connecting hardware, as generally used, is not designed as would be necessary in a framed structure, such as a truss, where movement in a joint would cause serious secondary stresses in the main members. Only one 3/4-in bolt is ordinarily employed, even in heavy H-frame construction in types of construction carrying wire heavier than has been general practice, and in very long spans the use of these connections, based entirely on experience, should not be followed without a careful check. The same applies to designs carrying heavy angles on crossarm construction where the entire angle load must be transmitted through the bolts to the pole. For such angles, the 3-pole structure is a more positive arrangement.

Conductor Arrangement and Spacing. In wood construction with short spans over comparatively level terrain, these parameters are determined largely by the line voltage. A wide variety of conductor arrangements is found in past practice. However, with the use of larger conductors and longer spans, the conductor configuration and separation are often a matter of providing the safest arrangement with ample spacing, especially for the occasional longer-than-normal spans encountered in rolling country. The conductor arrangement should provide spacing for these occasional long spans, as it is generally more economical to design a standard structure with spacings suitable for a span about 50% longer than normal rather than to use too many special structures.

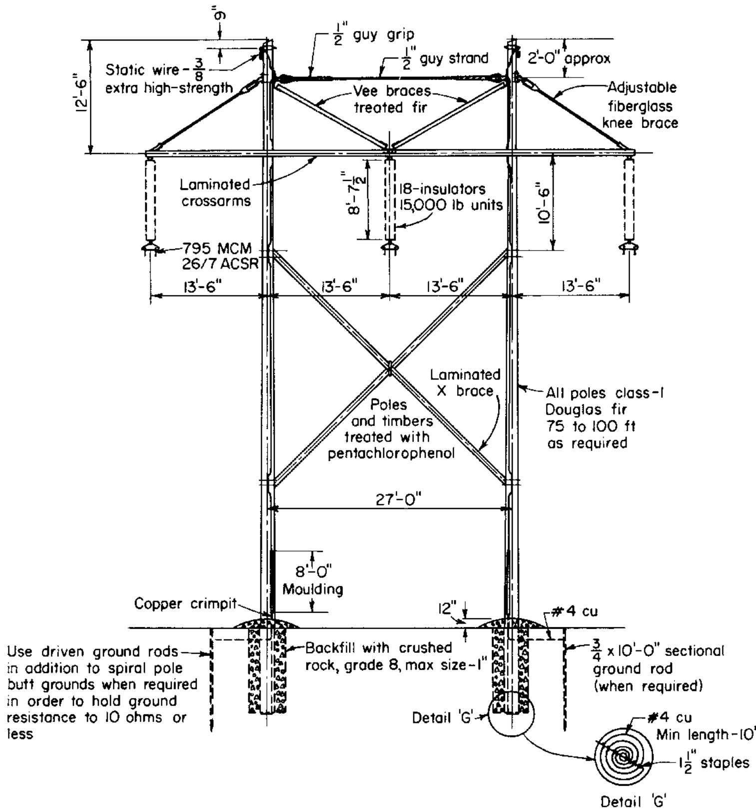


FIGURE 14-43 A 345-kV wood H-frame structure. (*Kansas Gas and Electric Co.*)

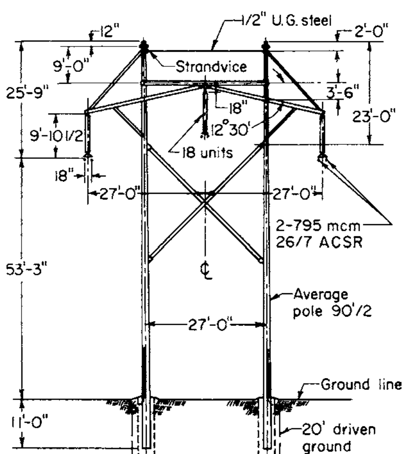


FIGURE 14-44 A 345-kV wood K-frame structure. (*Northern States Power Co.*)

The H-frame design gives one of the best conductor arrangements and mechanical strength for long-span construction and may be used as a special structure for especially long spans in almost any type of line.

Wood-pole H-frame structures are used on EHV lines at an appreciable saving over metal towers. Figure 14-43 shows an H-frame structure with trussed crossarm as used on the Kansas Gas and Electric Company 345-kV lines. This line uses two 795-kcmil ACSR conductors per phase on 18-in bundle spacing and 27-ft phase spacing. The insulator suspension hardware is not grounded, and full advantage is taken of the impulse insulation of the crossarm. Lightning flashovers would be expected to take place between the conductors and the ground wires on the poles and not to follow the insulator string and crossarm. All poles and timbers are penta-treated fir.

Figure 14-44 shows a modified H-frame wood-pole structure, designated a K frame, as used by the Northern States Power Company on its 345-kV system.[86] This

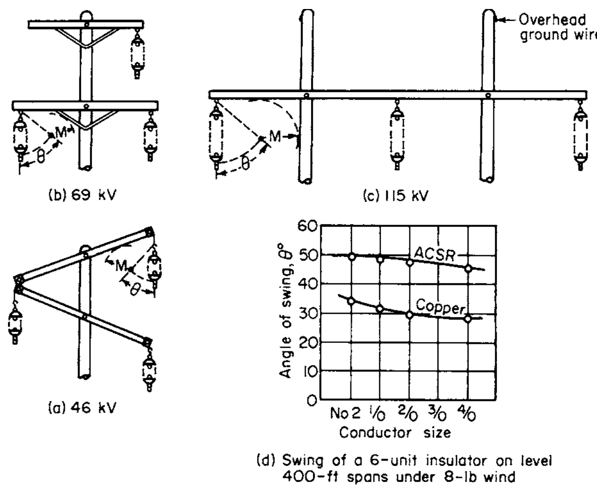


FIGURE 14-45 Typical suspension-insulator arrangements on wood construction.

structure also is designed to carry two 795-kcmil ACSR conductors per phase on 18-in bundle spacing and 27-ft horizontal phase spacing, but the center conductor is approximately 6 ft higher than the outside phases. This structure also takes full advantage of the impulse insulation of the crossarm.

It should be noted that the conductor spacing is primarily a function of the sag and that a conductor arrangement entirely satisfactory for a large or high-strength conductor would be hazardous for a small conductor, with correspondingly greater sags, in the same length of span. Also, a conductor spacing safe for a light loading district should not be used with the heavy sags required for heavy loading conditions. A small or lightweight conductor would require more spacing than a larger or heavier conductor in the same span and with the same sag.

On suspension construction the spacings are usually determined by the clearance required for swing of the suspension insulators as discussed under steel tower design. A detailed layout is required for each conductor, as the size and material have a marked effect on the swing characteristics (Fig. 14-45*c*). A fairly conservative assumption, which results in reasonable design, requires that the clearance from the conductor to a grounded structure shall be at least 0.75 the dry-flashover distance of the insulator or the "tight-string" distance under an 8-lb wind on the bare conductor at a temperature of 60°F. This may be modified in details, and it is common practice to allow somewhat reduced clearances to wood members. Typical layouts are shown in Fig. 14-45.

The swing should be taken for a somewhat more unfavorable case than level spans. The usual range of conditions encountered would be fairly well covered if a vertical span of three-fourths to two-thirds the horizontal span is assumed; that is, the clearances provide for cases where it is necessary to locate a structure somewhat below the elevation of the adjacent supports.

Angle structures in general use are shown in Figs. 14-46 and 14-47. The design is a matter of providing clearance from the conductor to the structure and to the guys under all conditions and at the same time of attaching guys as close to the load as possible to keep bending stresses down to a conservative value. On small angles where the loads are small, the angle pull may be carried as a bending in the pole and arm; but on larger angles, the loads should be carried directly by the guys, insofar as possible.

Figure 14-46*a* shows the usual small-angle construction, illustrating how it may be necessary to offset the arms to give clearance to the inside conductor. Figure 14-46*b* illustrates a similar design for small angles on heavy H-frame construction where the angle is so small that, if the maximum wind should blow from right to left in the illustration, it would cause the insulator to swing somewhat to the left of vertical. Therefore, clearance *M* must be provided, not only to the pole on the right, in the illustration, but also to the pole on the left. In the above designs the guy is attached some

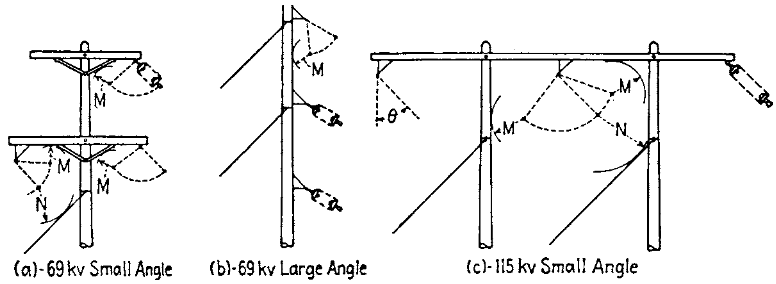


FIGURE 14-46 Suspension-angle structure.

distance below the crossarm in order to give a clearance N from the conductor to the guy, which is somewhat greater than the flashover distance across the insulator. The clearances N and M (Figs. 14-45 and 14-46) indicate the "normal clearance" and "minimum clearance," respectively. The normal clearance should be at least equal to the porcelain insulator.

The bracket on the pole as shown in Fig. 14-45*b* is used for larger angles where the mechanical stresses are too great for crossarm construction but for which the angle pull is not sufficient to swing the insulator string away from the pole under a wind from the right or at locations where a heavy vertical load is encountered on an angle structure. A similar 3-pole design is used for H-frame construction. The position of the insulator may be computed for various combinations of loading as shown below.

The simplest angle structure is illustrated in Fig. 14-46. The fewest pieces of hardware and the most direct transfer of stress to the guy are obtained. However, this design can be used only where the angle load is sufficient to hold the insulator string away from the pole under all conditions.

Angles greater than about 50° are usually dead-ended in a structure similar to Fig. 14-47, as it is not advisable to carry too large an angle on the usual suspension clamp. Erection is difficult on large conductors, and guying becomes complicated for very large angles on suspension construction.

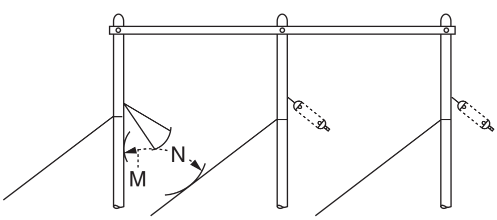


FIGURE 14-47 115-kV large-angle structure.

If grounded guys are used, a ground wire should be carried up the pole; and when the guy is attached close to the insulator, contact should be made with the insulator hardware to avoid the possibility of burning the pole from leakage or splintering from lightning. It is common practice to use one or two additional insulators on such angle structures.

Clearances on angles should be the same as on tangent construction, that is, under normal conditions must be somewhat greater than the flashover distance over the insulator but under maximum wind conditions may be reduced to 0.75 of normal, with some slight further reduction if this clearance is to ungrounded wood (see Table 14-23).

The greatest swing, that is, angle load and wind in same direction, may occur with wind on the bare wire at 0°F, but usually the combined ice and wind load is limiting because of the larger conductor tension. In the case of the wind blowing against the angle, clearances must be computed for a high temperature and resulting low conductor tension. Under normal conditions, for example, 60°F, full clearance should be maintained, equal at least to the dry-flashover distance of the insulator.

The angle load, that is, the transverse component of the conductor tension t at an angle α, is found as follows:

$$\text{Angle load} = 2t \sin\frac{\alpha}{2} \tag{14-84}$$

$$\tan\theta = \frac{(\text{angle load}) \pm (\text{wind load})}{(\text{vertical load}) + (1/2 \text{ weight of insulator})} \tag{14-85}$$

TABLE 14-23 Clearances for Various Lengths of Insulator String

Insulator, $5^3/_4$-in units	Normal clearance, in	Min. clearance 0.75 normal, in
4	$25^1/_4$	19
6	$36^3/_4$	27
8	$48^1/_4$	36
12	$71^1/_4$	50
16	$94^1/_4$	70

Note: 1 in = 25.4 mm.

in which θ = swing of the insulator from the vertical; the vertical load is the weight of the conductor supported by the insulator, or the weight per foot times the distance between the low points of the sag in the adjacent spans; and the horizontal load is the wind load on the spans supported by the insulator.

Pole Ground Wires. Ground wires should be installed on all poles, at all voltages, in lightning areas:

1. To prevent splitting of poles by lightning.
2. To provide a direct connection to ground and prevent pole burning if an insulator breaks down. Since the ground wire on these lines has relatively high resistance to ground, the wire can be small as No. 6 galvanized iron and the ground connection can be several wraps of the wire around the butt of the pole.

14.1.10 Line Accessories (Lines under EHV)

Suspension Clamps. These designs are fairly well standardized for the usual conductors. Simple, light, well-designed clamps in both malleable iron and forged steel are available for almost any conductor. The seat and clamping surfaces should be smooth, without any projections or sharp bends, and should be formed to support the conductor on long, easy curves and at the comparatively sharp bends formed at horizontal and vertical angles. Heavy, complicated clamps, unless very carefully designed, are generally avoided to allow as much freedom as possible at the support. For the same reason care is exercised to avoid rigid connections of any kind.

Trunnion-Type Clamps. These are designed to give an almost completely flexible connection by supporting the clamp on a pivot, approximately on the axis of the conductor (Fig. 14-48). Thus any vibration of the conductor tends to be transmitted through the clamp, eliminating much of the heavy binding stresses caused by a fixed support.

The suspension clamp is intended primarily to support the weight of the conductor and to prevent any longitudinal movement from accidental unequal tensions in adjacent spans. It is generally considered desirable but not always essential that the suspension clamp hold the conductor in case of a break. For large conductors under heavy tensions it is difficult to design a light, flexible connection that will not slip under such a contingency.

Slip, or Releasing, Clamps. Several especially heavy lines have been designed on the proposition that, since suspension clamps could not reasonably be secured that would positively hold the conductor, a clamp should be used that would hold under all ordinary conditions but would slip at something like one-half the maximum conductor tension in case of a break. This arrangement justified a considerable reduction in the exceedingly large longitudinal design loads on the towers and resulted in a considerable saving in tower and foundation costs. Several designs of slip clamps and releasing clamps have been used.

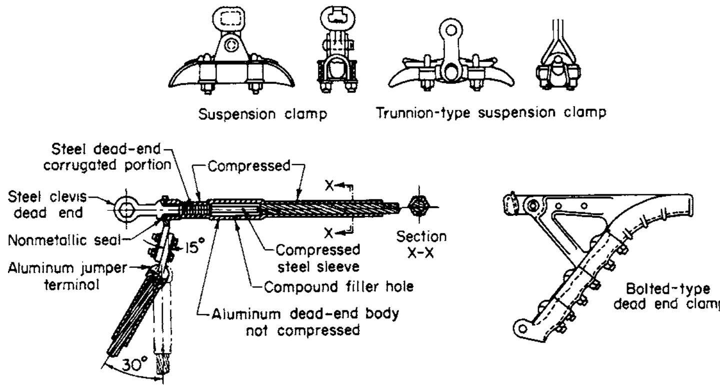


FIGURE 14-48 Conductor clamps.

Dead-End Clamps. These clamps are of the bolted type and are available for practically all copper and aluminum conductors. However, for the larger ACSR conductors the compression-type dead-end clamp is generally used (see Fig. 14-48). This is very similar to the compression splice used on ACSR.

The dead end for the steel core, which may have a clevis or an eye-type end, is pressed on after the aluminum sleeve has been slipped out over the conductor. The aluminum sleeve is then slipped back over the steel sleeve until the aluminum body makes contact with the shoulder of the steel sleeve. The electrical connection tongue on the aluminum body is aligned with the clevis or eye of the steel-core dead end as required, after which the aluminum body is filled with the nonoxidizing compound furnished with the body and the body is compressed. Similar pressed-on dead ends are available for copper, Copperweld, and other conductors. Several manufacturers furnish ACSR dead ends in all sizes required.

Armor Rods. These rods are quite generally used on ACSR lines as a protection against fatigue of the aluminum strands from vibration. Armor rods consist of a bundle of aluminum rods, somewhat larger in diameter than the strands of the conductor, laid parallel to the length of the conductor and arranged to form a complete covering. These are spirally twisted by a tool to lie approximately parallel with the lay of the strands in the cable and are clamped in place at each end. The suspension clamp is attached at the center, with the armor extending 2 or 3 ft on each side. The bending stresses caused by vibration are reduced by the increased diameter and area of metal and distributed over a longer section of conductor.

Vibration Dampers. The Stockbridge damper, as well as several other designs, are devices for damping vibration out of the entire span. Such dampers have been used on ACSR, copper, and steel conductors and ground wires, as illustrated in Fig. 14-49. The cause of conductor vibration and the action of the Stockbridge damper are outlined in the paragraph on wind-inclined motion.

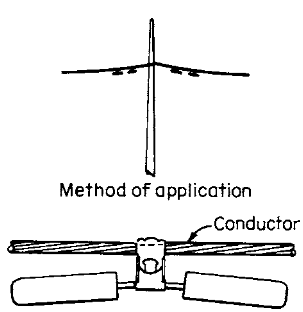


FIGURE 14-49 Vibration damper.

Overhead Ground-Wire Vibration. Overhead ground wires are especially subject to vibration; in fact, most steel ground wires will often be found in rather irregular vibration of small amplitude which generally does not appear to have any ill effects. Ground-wire attachments should be made with at least as great care as given the conductor clamps. Rigid clamps have been almost entirely abandoned in favor of a suspension clamp similar to that used on the conductor and attached by links or shackles so as to give a perfectly flexible connection.

Hardware. Many items of hardware have become fairly well standardized. The dimensions of the eye of eyebolts, the length of thread on various-length bolts, end links, and hardware for suspension insulators are quite uniform. It is usually possible to obtain about identical stock material from a number of manufacturers. Many other items such as shackles, guy clamps, and crossarm braces are furnished in such a wide variety that considerable care is required to choose the most commonly used but suitable stock items. Much expense and confusion in both construction and maintenance are saved by limiting the number of hardware items.

Insulating Braces and Guys. With the use of wood to increase the impulse insulation to decrease the line's sensitivity to lightning flashovers, steel crossarm braces have been replaced with wood on a number of lines. Connections are made by pressed-steel fittings. The use of a 48-in wood brace in place of steel, for additional wood insulation, is roughly the equivalent in lightning-flashover strength of adding one suspension unit to the insulation. The effect on 60-Hz flashover is, however, negligible.

To obtain equal wood insulation at guyed structures to what may be obtained on unguyed construction requires long wood insulators in the guys. These guy insulators are quite efficient because of the high tensile strength of clear wood; an ultimate strength of 6000 lb/in^2 on the net section is conservative. A 2- × 2-in fir insulator will develop the full strength of a 3/8-in Siemens-Martin guy strand. The design of the connection to the pressed-steel fitting requires only that several bolts of insufficient diameter be used to give the necessary bearing area between the wood and the shank of the bolt. The bolts should be placed alternately through the face and side of the stick to prevent splitting.

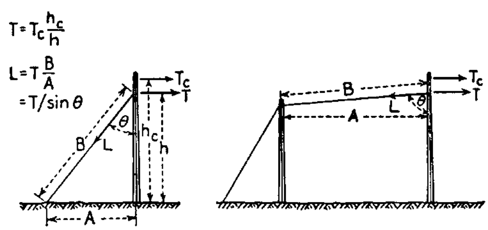
Reinforced fiberglass is receiving increased favor as guy-wire insulators in place of wood, and as crossarm braces in place of wood or steel. Impulse flashover voltages of fiberglass line hardware can be supplied by the manufacturers.

Guys. The various grades of guy strand are almost universally furnished in accordance with ASTM specifications. The ultimate strength for each size and grade is given in Sec. 4. The so-called double-galvanized is generally used. Common guy strand is not ordinarily employed in transmission construction, as the best-quality galvanizing is not furnished in this grade. Siemens-Martin strand is most commonly used for the lighter lines and high-strength strands for heavy construction.

More than one size of guy strand is not economical for a line, and often the same size may be used for several designs. The 3/8-in size, in either Siemens-Martin or high-strength grade, is most generally used both for guys and for overhead ground wires.

In the usual wood-pole construction great refinement is not required in designing guys. Usually it is sufficient to determine the number of guys, of the size and quality to be used on the line, required to support the load, an additional guy being employed for any fractional part. In transmission construction a safety factor of 2 is general for guys, although this may be somewhat reduced.

A common problem in guy design is illustrated in Fig. 14-50. The ratio of the guy load L to the conductor load T is the same as the



DESIGN DATA FOR GUYS

Guy	Ultimate strength, lb	Rod	Net area, sq in.	Ultimate strength, lb
3/8 in. s.m.	6950	5/8 in.	0.202	11,000
3/8 in. h.s.	10,800	3/4 in.	0.302	16,500
7/16 in. s.m.	9350	1 in.	0.551	30,500
7/16 in. h.s.	14,500			

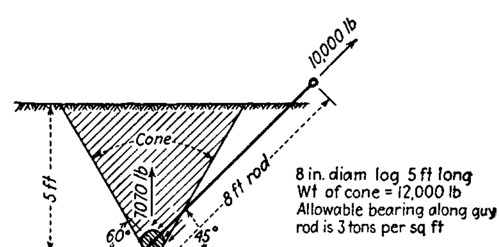


FIGURE 14-50 General guy and log anchor-strength calculations.

ratio of the length of the guy B to the distance A. The length B is readily determined from a sketch drawn to scale.

$$L = T\frac{B}{A} \qquad \text{or} \qquad L = \frac{T}{\sin\theta} \tag{14-86}$$

If the conductor load T_c is above the point of the attachment of the guy, then

$$T = \frac{T_c h_c}{h} \tag{14-87}$$

14.1.11 Conductor and Overhead Ground-Wire Installation

Conductor Stringing. This operation requires an experienced crew, not only to prevent damage to the conductor and overhead ground wire but also to maintain the sags and tensions specified in the design. Correct sags are essential to give the required mechanical safety, but it is equally important that the actual sag in the line correspond to that used in the design, to ensure proper clearances to the ground. More detailed procedures and guidelines are available in an IEEE publication.[87]

Stringing Equipment. All transmission conductors and overhead ground wires should be strung over free-running snatch blocks or rollers made for this purpose. Both conductors and ground wires of any material are easily damaged, and with the long spans and heavy conductors used in modern construction, satisfactory sags cannot be obtained at reasonable cost except by eliminating all possible friction at the supports. Dynamometers for measuring the tension are of value as a means of knowing the tension at all times, but they cannot be relied on to set the sag. The final sag should be adjusted by sighting. Walkie-talkie radios are widely used as standard erection equipment; better and more efficient work is obtained by having direct communication between the reel crew, the pulling crew, and the workers doing the sagging.

Sags are measured by setting sights on the structures at each end of the span at a vertical distance below the conductor support equal to the sag. This method is convenient and well within the necessary accuracy, even for inclined spans. For average inclined spans, the sag is taken as the sag for a level span of the same horizontal length, although the sag for a level span equal to the slope distance is more nearly correct. Except in extreme cases the horizontal and slope distances are practically the same. On long, inclined spans, when the low point of the sag falls below the ground level of the lower tower, it is more convenient to measure the vertical sag below the lower support.

Accuracy of Sagging. Friction in sagging blocks prevents the wire from reaching exactly the same tension at all points. As the wire is pulled up, the sag tends to be greater in spans from the pulling point; and when slacked back somewhat, more sag is thrown into the nearer spans. These effects are usually fairly well eliminated by allowing some time for the tension to equalize and by skillful handling.

Curves of "span versus sag" and "span versus tension" for possible stringing temperatures are used in sagging. The actual conductor temperature in sagging is very important. The IEEE Committee Report on Stringing and Sagging Conductors, *Trans. IEEE Power Group*, December 1964, p. 1235, recommends that the temperatures be obtained by direct measurement at the time of sagging by means of a thermometer placed inside a short length of the conductor and suspended at least 15 ft above the ground. Accurate temperature also is important in making allowance for creep during stringing. Creep elongation starts as soon as the conductor is pulled into the air, and it is important that this elongation be allowed to take place and not to pull the conductor back up to the calculated sag. One way to do this is to use a temperature curve which will indicate the calculated sag plus the additional sag due to creep up to the time of clipping in. The creep sag must be estimated from the manufacturer's curves.

In spans of varying length a greater sag tends to form in the long spans; and on steep grades the sags at the higher elevations tend to be less than at the bottom of the hill. These effects are not of importance except in extreme cases and are due to the fact that the wire is, and must be, supported on rollers in such a way as to be entirely free to travel. Thus the tension in the wire on each side of the roller must be equal irrespective of the slope of the wire away from this support, and the resultant on the support is not vertical but in the direction of the bisector of the angle between the slopes of the wires as they leave the roller on each side. At a support between a short and long span (Fig. 14-51) the wire on the short-span side is more nearly horizontal, *OA*, whereas on the long-span side the wire may have a considerable slope *OB*. The tension on each side must be equal, $OA = OB$, but the horizontal component of the tension is therefore less in the long span *BD* than in the short span *AE*. The resultant *OC* is inclined.

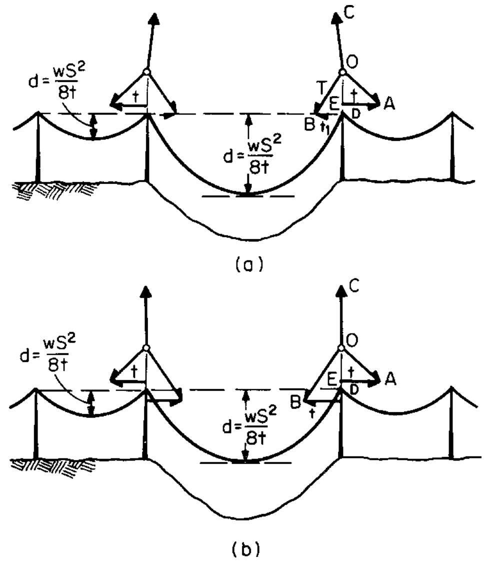


FIGURE 14-51 Diagrams illustrating the change in tension in long spans: (*a*) on rollers; (*b*) clamped in theoretically correct position.

It is theoretically possible, although very difficult practically, to clamp the conductor at the correct position so that the resultant will be vertical and the horizontal tensions equal as in Fig. 14-51. This is the condition assumed in the computations and office location and, for all except extreme cases, is the reasonable assumption.

Similarly, the different slopes of the wire leaving the roller on hillsides with spans of equal length but at different elevations cause the horizontal component of the tension to be less in the wire with the greatest slope (Fig. 14-52). With a series of spans on a slope this effect tends to accumulate, for the horizontal tension t_2 at the upper support must be the same as the horizontal tension t_2 at the lower support, whereas the resultant tension T_2 is less than the tension T_1 because of its smaller slope.

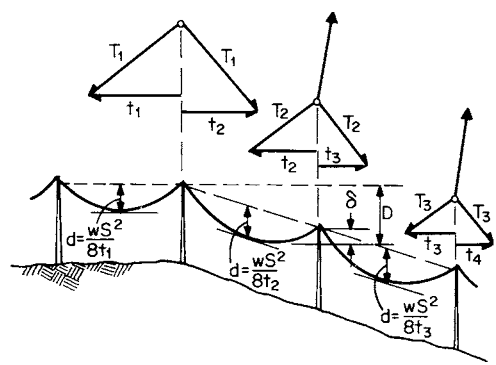


FIGURE 14-52 A much-exaggerated illustration showing the change in tension on hillsides.

The differences in sag are not usually carried from one conductor pull to the next, but each pull is sagged to approximately the correct tension independent of the other; thus when the snubs between pulling sections are removed, differences in tension tend to equalize. For this reason it is best not to clamp in the conductors too close behind the sagging crew. Often skillful sagging reduces these effects by using the friction in the blocks to prevent the conductor from "collecting in the low spots."

These irregularities are of little consequence generally, especially when it is realized that the important consideration is to have equal tensions under maximum load rather than under bare-wire conditions. In extreme cases, provision for the above conditions may be made by special sags, allowing somewhat higher tensions in spans above the normal level of the line and providing extra clearance in low sections so that slightly larger than normal sags may be used. Occasionally special methods must be devised.

Compression Joints. These joints are used with the large ACSR conductors and the large "all-aluminum" conductors, including the high-strength types. As with the compression dead ends, the most widely used joints are those made by Somerset Products Co., a subsidiary of Thomas & Betts Corp., and the Alcoa Conductor Products Co. of the Aluminum Company of America. Cutaway drawings of these joints are shown in Fig. 14-53. They consist of aluminum sleeves and steel sleeves for ACSR. Installation procedures call for the aluminum cable and the insides of the aluminum sleeves to be thoroughly cleaned.

If the conductor is weathered, the strands should be unlayed and all scale removed. The aluminum sleeve is then slipped on the cable and backed out of the way. The aluminum on each cable end is next carefully cut back, care being taken not to nick the steel core, for a distance equal to one-half the length of the steel sleeve plus a distance of 1/2 in or more, depending on the size of the conductor, so that the elongation of the steel sleeve on compression will not interfere with the free lay of the aluminum strands. The conductor ends are then marked by tape or other suitable means to center the sleeve. The steel sleeve is put in place and compressed, working from the center out. The aluminum sleeve is next slipped into place and filled with the nonoxidizing compound furnished with it, the filler holes are plugged, and the joint is ready for compression. The sleeves are compressed by working from the center out. The center section of the aluminum sleeve over the steel sleeve is not compressed. When the compression is completed, the Alcoa sleeve is hexagonal and smooth from overlapping compressions, and the Thomas & Betts sleeve, also hexagonal, has uncompressed ribs between the compressions as shown in Fig. 14-53.

Overhead Ground-Wire Installation. Overhead ground wires should receive no less care in erecting than the conductors because the usual zinc or copper protective coating is very easily destroyed. Ground wires should be sagged in the same way as the conductors except that the important factor in ground-wire sags is to maintain ample clearance to the conductors. Generally, ground wires are sagged to about 80% of the conductor sags, thus ensuring proper clearance even under ice loads. McIntyre joints are frequently used for splicing, but a much more efficient joint can be made with the pressed-steel joint similar to the ACSR joint illustrated in Fig. 14-53.

In addition to wind and ice loads, overhead ground wires are subject to lightning damage and burning at connection points due to power frequency fault currents. Ground wires are sometimes equipped with jumper connectors at tower attachment points to help bypass fault currents. Three-eighths-inch steel ground wire is reasonably resistant to most lightning currents, but positive lightning strokes, or "winter lightning," although relatively rare, can heat the stricken point sufficiently to damage the wire or possibly to cause it to separate.

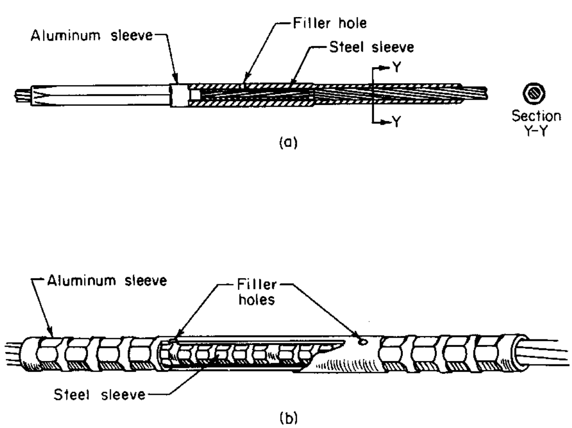


FIGURE 14-53 Compression joints: (*a*) Aluminum Company of America (*b*) Thomas & Betts.

14.1.12 Transpositions

Transpositions[88,89] are made for the purpose of reducing the electrostatic and electromagnetic unbalance among the phases which can result in unequal phase voltages and currents for long lines. Untransposed lines also can cause inductive interference with paralleling wire communication lines. However, communication interference in the past has been largely with overhead long-distance telephone and telegraph lines. Most of these lines are now going underground, and other overhead lines are being replaced by microwave radio.

For some time, transpositions were little used. With the large power-system networks comprising most of the country's transmission lines, the unbalance of an untransposed line is largely smoothed out by the phase-balancing effect of the rotating equipment scattered over the system. The transpositions, however, do enhance the reliability and efficiency of the line in the following respects: (1) they restrain the amount of current which one line will induce in the other, thereby enhancing the reliability of fault arc extinction in the event of a faulty circuit; (2) they serve to reduce transmission power losses; (3) depending on their location, transpositions can serve to reduce the electromagnetic coupling of power-line currents in adjacent telecommunication lines.

14.1.13 Operation and Maintenance

Operation. Effective operation of a system is as essential to good service as is excellent engineering design. In fact, a well-designed system may fall short of its service requirements owing to faulty operation. Aside from switching lines and power units to meet the load conditions of the system, operation consists not only in restoring service promptly after an interruption but also in detecting and removing faulty apparatus, thus actually preventing the development of faults.

A chief system operator should be in absolute control of the system, and if it is a small system, he or she should have direct communication with and direct control over every part of it. If it is a large system, it will not be possible for one person to supervise all switching operations, and area dispatchers must be located at convenient points. These dispatchers will have the same authority over their areas as the chief system operators for small systems. The area dispatchers will call on the chief system operator not only for approval of unusual switching operations, particularly those involving interruptions to important loads. They will, however, make reports each shift, as convenient, on routine switching operations and will report major interruptions as soon as possible, with cause if known. Dummy boards are useful at dispatching centers. These boards should show the one-line diagrams of the circuits at all stations under the dispatcher's supervision, and provision should be made to show whether switches are open or closed. These boards must be kept correct up to the minute, even if it is necessary to do so by temporary means. It is best to anticipate system changes so that the dummy board can show the changes as soon as they are made. During normal operating conditions no switching should be done, including that of generators, without the dispatcher's permission. All dispatching orders should be reported back in order to prevent misunderstandings and should be recorded in log books both by the dispatcher and by the operator who will do the switching. The logs should show a record of all transactions, with particular care about times of receiving orders, of opening and closing switches, and of cases of trouble.

Emergency routines should be set up for all stations and should be followed at times of catastrophic storms when all means of communication with the dispatcher are interrupted. These routines will list the sequence for doing emergency switching on the operator's responsibility in an effort to restore service.

Supervisory control systems make it possible for operators at one transmission substation to operate several nearby substations as well as their own and thereby reduce operating personnel. Supervisory control also makes it possible for one central dispatching office to operate all the transmission substations serving a large metropolitan area. The supervisor may utilize carrier-current, fiber-optic wires, microwave radio, or telephone channels for transmitting information and operating switches.

Sleet or glaze formation on lines is highly undesirable, and some companies prevent it by raising the temperature of the conductors with current. Ice will form on conductors over a small temperature range which is on the order of −3 to +2°C. The current required to prevent ice formation may be found according to Clem[90] from

$$I^2 = \frac{\theta\sqrt{dv}}{8.18 \times R} \times 10^4 \tag{14-88}$$

in which I = current, θ = temperature rise in degrees Celsius above surrounding air, R = conductor resistance in ohms per mile at 20°C, d = diameter of conductor, and v = wind velocity in miles per hour.

With the lines in service, it is usually difficult to obtain sufficient current. However, the necessary current sometimes may be obtained by transferring load from other lines to the line in trouble. Dead lines may be heated by short-circuiting them at one end and sending the necessary current from the other end. The approximate voltage to neutral is $E = I\sqrt{R^2 + X^2}$, where R = resistance per wire and X = reactance per wire.

Melting the ice after it has formed is considerably more difficult and requires more current than is required to prevent formation. Clem's article gives the various formulas required to calculate the melting current.

Maintenance. Periodic inspection is normally maintained over all lines, with the frequency of inspection depending on the country traversed and the importance of the lines. In some densely settled areas, patrols of once a week may be considered necessary, whereas important lines in areas subject to heavy storms or other hazards may not require inspections more than once in 2 months. Patrollers may cover the line on foot, on horseback, by automobile, by all-terrain vehicles (ATVs), or by helicopter, depending on the characteristics of the right-of-way. Close and accurate patrolling is not obtained by one person in an automobile, in general, even when the line follows a highway. Helicopter patrol is by far the best over mountainous and sparsely settled country, and 200 mi a day can be covered readily. The helicopter can fly as slowly and as close to the line as is necessary, and it has the great advantage that the patroller is looking down on the line instead of up against the bright sky as a background. Tower and wood-pole structure numbers should be fastened to the tops of the towers or structures in such a manner that they can be read without trouble by the person in the helicopter. Helicopter patrol cannot be used over urban areas or congested industrial areas because of governmental restrictions on height of flight. Patrols on foot are best in such areas. Horseback patrol is best in cattle-range country, if aerial patrol is not available.

Landslides, washouts, danger timber, or anything else that is a potential danger to the line, such as piles of brush or straw which if burned could cause hot gases to short circuit the line, should be reported. Of course the person on patrol must also be on a close lookout for damaged conductors, insulators, and structures. A pair of field glasses is usually considered indispensable.

Personnel on ground patrols should keep the dispatcher informed as to their whereabouts and should call in from all patrol telephone stations and from substations as they reach them. They should call in not less often than morning, noon, and night. If a storm comes up while patrollers are out on a line, they should call the dispatcher as soon as possible, telling where they can be reached. Patrol cars and helicopters are radio-equipped.

Tree gangs whose sole duty is to remove brush, trim trees, and remove danger timber have been found to be advantageous by large companies. The use of chemical sprays to kill brush along rights-of-way is satisfactory from the standpoint of killing the brush if legally permissible but leaves a potential fire hazard. Care must be taken in spraying that wind does not blow the chemicals over growing crops. This danger has been found to be a disadvantage in helicopter spraying.

Emergency crews are stationed at locations always available by telephone or radio so that every important section of the transmission system can be reached by a crew within a reasonably short time. Small houses containing spare parts, such as insulators, lengths of cables, and clamps should be located at intermediate and accessible points along the line in sparsely settled country. Such houses should be kept locked. Some companies employ concrete construction with iron doors. A routine inspection and checking of materials in such houses is advisable. A light truck, provided with two-way radio and with the necessary tools and materials for making immediate repairs, is used by many companies. In addition to spotlights on the truck, a spotlight operated from a portable storage battery is frequently very useful. Small houses containing spare parts, such as insulators, lengths of cables, and clamps, should be located at intermediate and accessible points along the line in sparsely settled country. Such houses should be kept locked. Some companies employ concrete construction with iron doors. A routine inspection and checking of materials in such houses is advisable.

Line repairs and replacements are accomplished either when the line is energized or deenergized. The line crew should notify the dispatching office when a particular line or section is desired. If deenergized maintenance is desired, the line should then be not merely cleared through the circuit breakers but opened by disconnecting switches as well. If it is to be out of service for an extended period and there is danger of lightning, the line should be grounded out at its terminals to prevent the possibility of flashing over switches or insulators at the terminals. If the line is not equipped with grounding switches, it may be grounded out by equipment such as is used by line crews. Before the line crew is allowed to work, the line should be short-circuited and well grounded on each side of the location where the crew will work, with the grounding equipment in full sight. The grounding equipment should consist of heavy extraflexible copper cables, which should be attached by means of "hot-line" tools and clamps, the line being considered to be "hot" until the grounding equipment is applied. Ground chains are not safe and should not be used. Reliance should not be placed in grounding switches or grounding cables at the ends of the line.

In order to make repairs and replacements without interrupting the service, special "live-line" tools have been devised whereby insulators may be replaced, conductors spliced, etc., on lines of all voltages while the line is hot. Live-line maintenance methods are described in an IEEE Task Group Report[91] of the IEEE Transmission and Distribution Committee (1973). It covers methods and equipment for live-line maintenance. A foundation is presented from which working clearances and methods can be developed for specific needs in particular applications. A bibliography[92] concerning live-line maintenance is available.

Damaged insulators and insulator sections may be detected while they are in service, but suitable precautions should be followed. In general, the methods employed for faulty-insulator detection are based on the measurement of the voltage gradient across the individual units of a string of suspension insulators or across the parts of multipart pin-type insulators. For safety reasons, none of the test methods should be used in wet weather.

Faulty insulators may also be detected by special radio interference locators consisting essentially of a sensitive battery-operated receiver coupled to either a directional loop antenna or a "whip" antenna. The latter type may be attached to a hot-line stick to enable close investigation of the insulator under test. Infrared techniques are also employed.

The Doble method[93] is, in effect, a spark-gap voltmeter which is safe to use and which gives high accuracy in measuring potentials in the field on live transmission-line insulators.

There are two general types of Doble safety tester: the type A single-prong tester for multipart pin-type insulators on either wood or steel construction, and the type B two-prong tester for multiunit suspension-type insulators on either wood or steel construction. The type B is most applicable to transmission lines.

The equipment consists of a micrometer spark gap in series with a capacitor and a special telephone-type headset with which to listen to gap sparkover. The telephone receiver is heard through a rubber hose connected to a highly insulated hollow tube which is long enough so that there is no danger to the operator.

Both sides of the electric circuit of the tester are terminated by exposed metal tips (Fig. 14-54) arranged so as to bridge readily a single insulator disk or section. In the circuit between the two contact tips, a protective insulating capacitor is built into the tester in such a manner as to make the impedance between its terminals greater than the impedance of a single good disk. Thus, in operation, the tester *does not short-circuit* the disk under test. The tester may be considered as a voltmeter which indicates the voltage between the points on the insulator touched by the tips of the tester.

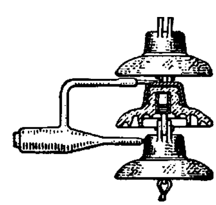


FIGURE 14-54 Doble insulator tester using the principle of a spark-gap voltmeter for suspension insulators.

The degree of defect in the disk under test is indicated by the size of the maximum gap at which a noise is heard in the tester, as compared with the size of a maximum gap for a good disk in the same position in a string; a

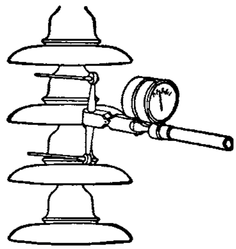


FIGURE 14-55 The I-T-E Imperical Corp. two-prong live-line insulator tester.

totally dead disk gives no sound, irrespective of the gap setting. In practice, one gap length is fixed in advance and used for all units of the string. The operator judges if a unit is defective from the noise heard in the headset. This apparatus is in use for testing insulators on lines at voltages from 11 kV through the medium transmission ranges.

The I-T-E Imperial Corp. live-line insulator tester detects defective insulators, either multipart pin-type or suspension, by comparing the measured voltage distribution over insulators or insulator strings while in service with characteristic curves plotted for good insulators of the same type. It is suitable for use in testing insulator integrity on transmission systems with nominal voltages through 230 kV. The tester employs a single-prong head for multipart pin-type insulators, a small two-prong head (Fig. 14-55) for suspension strings or small one-piece insulators, and a large two-prong head for multipart pedestal insulators. Visual indication is given by means of a meter which shows a deflection in proportions to the voltage gradient. Since relative indications only are required, the meter is calibrated simply in units of deflection. Tests may be made of all shells of multipart pin-type insulators and all units of suspension strings with equal facility. As with the Doble tester, tests should be made only on perfectly dry insulators.

Doble field power-factor test is used (in contrast to that previously described) for testing the insulation of electrical power apparatus *with the apparatus out of service*. Dielectric watts loss and charging current are measured at selected test voltages up to 20 kV, from which power factor, capacitance, ac resistance, and the presence of ionization (corona) can be determined. The specimens to be tested may be in the substation or in the service building.

The test equipment is capable of determining the condition of electrical insulation of bushings, bus supports, cables, capacitors, circuit breakers, insulators, surge arresters, liquid insulation, potheads, rotating machinery, transformers, and voltage regulators. Power-factor measurements with this equipment have been adopted by many companies as a criterion for servicing power-apparatus insulation.

High power factors or sudden increases in power factor from a previous test indicate contaminated or deteriorated insulation which may be an operating hazard. Changes in the watts loss, charging current, ac resistance, and capacitance between tests are also used for indicating operating hazards in apparatus insulation.

A variety of other makes of portable insulation testers, both ac and dc, are available for use in testing line insulation, if desired, when the line is out of service.

14.1.14 Foundations

Lattice-Tower Foundation Loads and Displacement Criteria. Lattice-tower foundation loads consist of vertical tension (uplift) or compression forces and horizontal shear forces. For tangent and small-line-angle towers, the vertical loads on a foundation may be either uplift or compression. For terminal and line angle towers, the foundations on one side may always be loaded in uplift while the other side may always be loaded in compression. The distribution of horizontal forces between the foundations of a lattice tower vary with the bracing of the structure. A typical free-body diagram of foundation loads is shown in Fig. 14-56.

When the foundations of a tower displace and the geometric relationship of the tower to its foundations remains the same, any increase in load due to this displacement will have a minimal effect on the tower and its foundation. However, foundation movements which change the geometric relationship between the tower and its foundations will redistribute the loads in the tower members and foundations. This will usually cause greater reactions on the foundation that moves least relative to the tower, which in turn will tend to equalize this differential displacement.

Presently, the effects of differential foundation movements are normally not included in tower design. Several options are available should the engineer decide to consider differential foundation displacements in the tower design. These options include designing the foundations to satisfy performance criteria which will not cause significant secondary loads on the tower, or design the tower to withstand specified differential foundation movements.

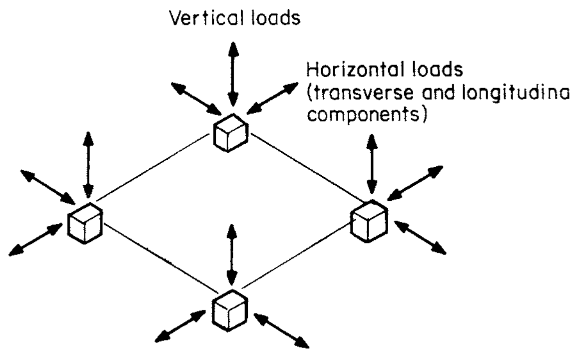


FIGURE 14-56 Typical loads acting on lattice-tower foundations.

Single-Shaft Foundation Loads and Displacement Criteria. Single-pole structures have one foundation so that differential foundation movement is precluded. The foundation reactions consist of a large overturning moment and usually relatively small horizontal, vertical, or torsional loads. Figure 14-57 presents a free-body diagram of the loads.

For single-shaft structures, the foundation movement of concern is the angular rotation of the shaft in the vertical plane and horizontal displacement of the top of the foundation. When these displacements have been determined, the displacement of the conductors can be computed. Under high-wind loading, a corresponding deflection of the conductors perpendicular to the transmission line can be permitted. Accordingly, a large ground-line displacement of the foundation could also be permitted. As a result of foundation rotation, the clearance between the conductors and the structure would be decreased only for structures with single-string insulators. The midspan ground clearance and the change in line angle would also decrease a negligible amount.

In establishing displacement criteria for single-shaft-structure foundations, consideration should be given to how much total, as well as permanent, displacement can be permitted. In some cases, large permanent displacements might be aesthetically unacceptable and replumbing of the structures and/or their foundations may be required. In establishing displacement criteria, the cost of replumbing should be compared to the cost of a foundation that is more resistant to displacement.

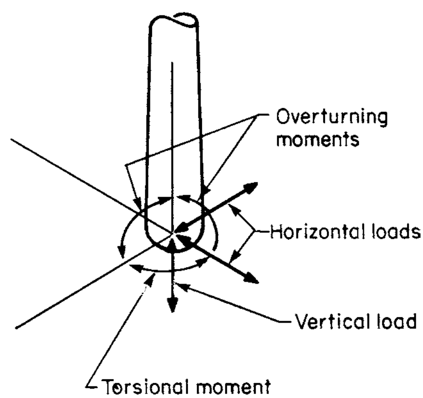


FIGURE 14-57 Typical loads acting on foundations for single-shaft structures.

For terminal and large-line-angle structures, large foundation deflections parallel to the conductor may be intolerable. For these structures, excessive deflections may reduce the conductor-to-ground clearance or affect the load capacity of adjacent structures. There are also problems in the stringing and sagging of conductors if the deflections are excessive. These problems are usually resolved by designing a more deflection-resistant foundation, construction methods, or use of permanent guys.

Framed Structure Foundation Loads and Displacement Criteria. These structures are dependent in part for their stability on one or more of their joints resisting moment. The foundation reactions are dependent on which joints can resist moment and the relative stiffness of the members. Figures 14-58 and 14-59 present free-body diagrams of four- and two-legged framed structures. If the bases of

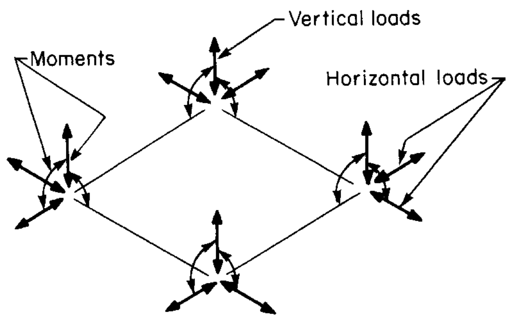


FIGURE 14-58 Typical loads acting on foundations for four-legged framed structures.

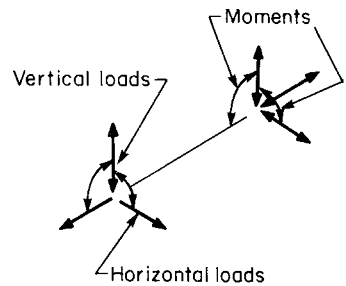


FIGURE 14-59 Typical loads acting on foundations for two-legged framed structures.

structures are assumed as pins or universal joints, then the moments acting on the foundations will theoretically be zero.

Many different types of two-legged, H-framed structures are in use in transmission lines. This has been particularly true since visual impact has become of greater concern.

The H-frame structure is particularly applicable for wood, tubular steel, and concrete poles. The crossarm may be pin-connected to the poles. These structures may be unbraced, braced, or internally guyed as shown in Fig. 14-60.

As with lattice towers, past practice has not usually included the influence of foundation displacement and rotation in H-frame structure design. Significant foundation movements will redistribute the frame and foundation loads. The foundations can be designed to experience movements which will not produce significant secondary stresses, or the structure can be designed to a predetermined maximum allowable foundation displacement and rotation.

Typical HV and EHV Guyed Structures. *Guyed Portal Structures.* The design and utilization of guyed-mast transmission structures (guyed towers) has evolved very rapidly since the mid-1960s or so, and the subject warrants discussion. A high-strength wire stranding used in tension and a latticed mast in compression with limited bending are two of the most efficient structural components, a point not lost on designers of transmission lines in northern Europe during the early days of transmission of electricity.

The most popular form was the guyed portal tower (Fig. 14-61), in which two tripods, each essentially consisting of a mast and two guys, were held apart at the tops by a rigid crossarm. The arrangement permitted the use of two mast footings (compression with a little shear load) and if the structure was not too tall, the four guys could be brought to two anchors, each doing double duty. The structure also offered the possibility of easy ground assembly and erection by one-piece rotation using a gin pole or A frame. Portal tower application was to a large extent limited to flat terrain and relatively short towers (short spans or low voltages) as rough terrain required masts of different lengths, which

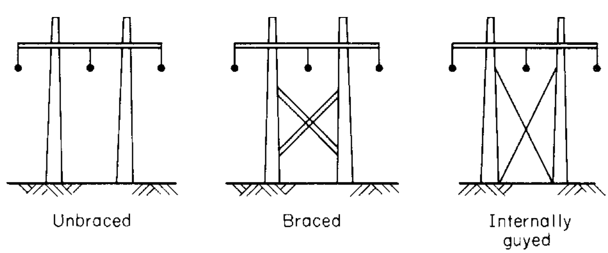


FIGURE 14-60 Typical H-frame structures.

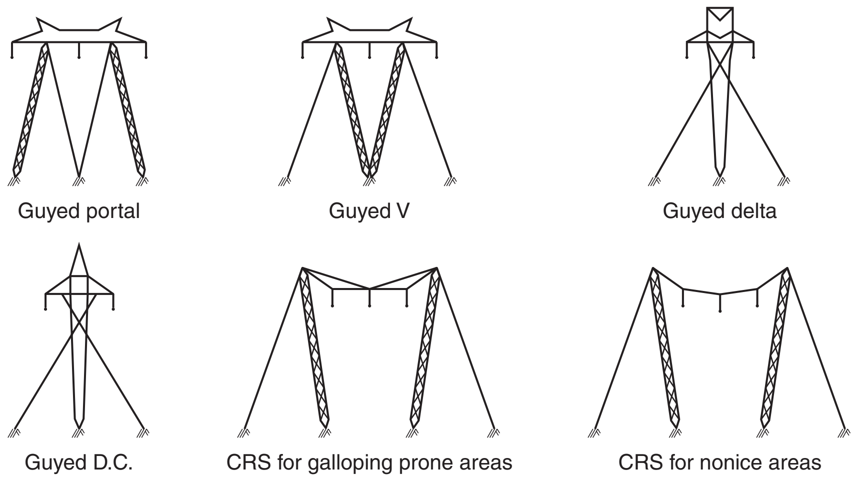


FIGURE 14-61 Typical externally guyed structures.

complicated the usual erection methods and tall structures made common guy anchors an impossibility or else greatly increased the tower stresses under transverse loading.

Guyed-V Tower. In 1957, the guyed V tower was developed[94] so that the guyed-mast principle would be more suitable for rough terrain; the guyed V consisted of the same two tripods moved laterally but was now held together by the crossarm.

Introduction of the guyed V helped promote the initial and serious use of helicopters for line construction for the first guyed-V towers, which were designed in aluminum. The very light weight of both the aluminum material and guyed-mast concept permitted transport and erection of a complete structure with the limited capacity of the helicopters then commercially available. Guyed-V tower use expanded rapidly with thousands of kilometers of lines built at up to the 500-kV level, fabricated in both steel and aluminum as the economies of the V principle were proving to be sufficient to justify use even when helicopter erection was not warranted.

Application expanded into the 750-kV class in the United States, Brazil, and Canada, and nominal spans were consistently in the 450- to 500-m range as the low cost of the extra height of taller towers (two masts and some guy wire) were extending the optimum envelope.

Single-mast guyed towers such as the delta were developed that offered compaction benefits and were also easy to use on rough terrain as they also require only one compression footing; the guy lengths are cut as required to fit the terrain. The guyed-V tower became the most widely used tower at the higher voltages, although all forms of guyed-mast structures were usually restricted to open terrain and areas where widespread guys were not considered to be hazards to the use of large farm equipment or where land occupancy was not a problem.

Cross-Rope Suspension (CRS) or Chainette. The next step in guyed-mast development followed from the failure of a 750-kV class tower due to material defect, a failure in a remote area where construction had been by a large mobile crane that could move only on the frozen winter roads. The replacement guyed-V structure was too heavy for available helicopter lift, and the repair was greatly delayed until crane access came with winter weather.

As voltages increase, all towers tend to become top-heavy, and the guyed-V or guyed-Δ are no exceptions. The urge to dispense with the crossarm led to the development of the *cross-rope suspension* (CRS) tower or *Chainette*, as it is referred to in Quebec, Canada.

The CRS concept was actually derived from the successful CRS system built in the mountains of British Columbia, Canada in 1955[95,96] and uses one or more wire ropes suspended between two guyed masts to replace the crossarm and support the insulator strings. Even at 1000 kV, the individual masts are well within the capacity of modest low cost helicopters, and thus the problems of initial construction and emergency replacement are easily solved, with crane or helicopter. The initial design of the CRS[97,98] made use of the six-part suspension assembly because it was feared that galloping of a single phase might find the support point forces transferred through the rope system to the other phases and thus promote widespread and damaging activity. The CRS found application on the third, fourth, and fifth lines of the 735-kV James Bay system in Quebec and on a section of 500-kV lines in Oregon, in the United States.

The next applications and development have come about in South Africa at 400 kV and in the Argentine at 500 kV, where a single cross rope is used because of the absence of any significant icing and thus no fear of galloping.[96]

All CRS lines must use a special construction or spacer rope extending between the tops of the masts; the initial use is to position the mast tops before the conductors are in place to provide tension to the system and subsequently to provide means of access to the phases by use of a wheeled ladder for both stringing and sagging work and for maintenance.

The CRS type of construction has a few negative aspects as it requires space at the tower sites (an open area that can be farmed), but the positive points are many. Both single- or multiple-rope CRS systems permit reduced phase spacing with increased SIL values and limited only by gradient effects or fear of wind clashing. On the 1150 km of the 500kV line in the Argentine studies showed that clashings were a negligible threat and the reduced phase spacings available by the CRS increased SIL values and reduced compensation costs by many millions of dollars. All wire rope components including the guys are precut to length and can be tested to working loads, and length adjustment is included in only one of the guys. The weight of structural steel in a tower will be about 50% of that of a comparable guyed-V or guyed-Δ tower. With overhead wires attached directly to the tops of the masts, the lightning protection is about as good as possible and the extra relaxation provided by the suspended wire assembly added to the insulator strings ensures that cascading potential is negligible.

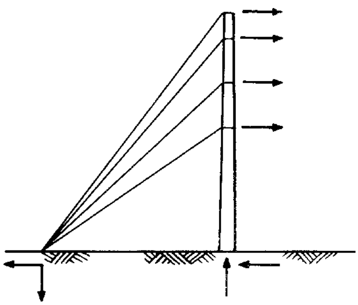


FIGURE 14-62 Single-shaft externally guyed structure.

Guying Rigid Structures. The guyed structures of the preceding section consisted of structural components pivoted or free to rotate at the connections to the foundations. Thus, the guy loads are in all cases determined by simple static analysis. The addition of guys to structures that are fixed or rigidly attached to the foundations (Fig. 14-62) produce arrangements that are indeterminate to varying degrees and for some, the analysis of strength can be quite complex and performance will be dependent on introducing and maintaining precise levels of guy pretension.

The problems arise because of the large strains or stretch that guys develop in order to resist the loads and the varying degrees of rigidity of the structures that they are guying. Guyed-pole structures, as shown in Fig. 14-62, if they are wood and single-pole or the very common H-frame structure, are readily guyed, since the wood components are relatively flexible and allow the guys to develop load before the poles deflect enough to fail.

However, guyed-metal-pole structures that are fixed at their bases become more difficult to assess and analysis is usually needed to confirm both the distribution of loads between the guys and the poles themselves and also to set the pretension needed in the guys to ensure that the deflections or distortions of the components remain within limits. The design engineer must be conscious of the fact that deflections of poles can introduce large additional bending stresses caused by the P-Δ effect—the pole compression acting on the moment arm of the bow of the pole.

The most complex situation arises from attempts to guy rigid lattice structures (Fig. 14-63). In order to resist a load applied to the structure, the guys will normally stretch so much that the rigid structure will already have failed. Successful guying in such a manner requires either oversized guys to limit the stretch of the guys or very precise levels of pretension, set and maintained to distribute the loads in desired manner between the structure and the guys.

There is one condition under which guying of a rigidly framed structure is readily done and that is when the purpose is to restrict the movement of the structure on failure. Longitudinal guying of rigid structures or even of H frames is a method of creating stop towers to limit the movement of slack and thus stop a cascading of structure if one is under way. Transverse guys are sometimes applied outside of heavy angle structures to restrict their failure mode and thus limit the amount of slack that can be introduced into the line system if the tower failure occurs under an extreme ice loading.

Foundation Types. A wide variety of foundation types can be used with self-supporting or guyed lattice, framed, and single-shaft structures. They include the following:

Lattice tower	Framed and single-shaft structures
Steel grillages	Concrete spread foundations
Concrete spread foundation	Drilled shafts
Rock foundation	Direct embedment
Drilled shafts	

Steel Grillages. Figure 14-64 shows three typical types of steel grillages. Figure 14-64*a* is a pyramid arrangement in which the leg stub is connected to four smaller stubs which in turn are connected to the grillage at the base. The advantage of this type of construction is that the pyramid can transfer the horizontal shear load down to the grillage base by truss action. However, the pyramid arrangement does not permit much flexibility for adjusting the assembly, if needed. In addition, it is difficult to compact the backfill inside the pyramid.

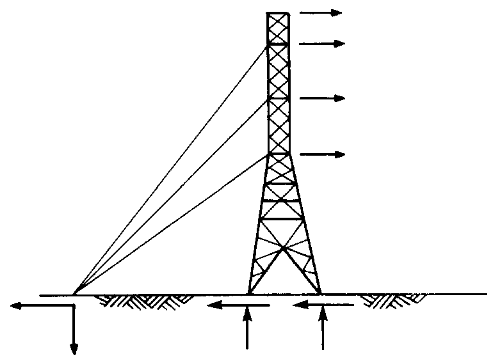


FIGURE 14-63 Externally guyed lattice tower.

Figure 14-64*b* shows a grillage foundation which has the single leg stub carried directly to the grillage base. The horizontal shear is transferred through shear members that engage the passive lateral resistance of the adjacent compacted soil.

Figure 14-64*c* also has the single leg stub carried directly to the grillage base. This type of grillage foundation has a leg reinforcer which increases the area for mobilizing passive soil pressure as well as increasing the leg strength. The shear is transferred to the soil via the leg and reinforcer and resisted by passive soil pressure. The base grillage of these three typical foundations consists of steel beams, angles, or channels which transfer the compressive or uplift load to the soil.

The advantages of steel grillage foundations are that they can be purchased with the tower steel and concrete is not required at the site. The disadvantage is that these foundations usually must be designed before any soil borings are obtained and may have to be enlarged by pouring a concrete base around the grillage if actual soil conditions are not as good as those assumed in the original design. In addition, large grillages are difficult to set with required accuracy. The placement and compaction of the backfill material are critical to the actual load-carrying capacity and load-displacement characteristics of the foundation.

Concrete Spread Foundations. This type of foundation consists of a base mat and a square or round pier. It is constructed of reinforced concrete. There are several variations as indicated in Fig. 14-65. The stub angle can be bent and the pier and mat centered. The mat can be located so that the projection from the stub angle intersects the centroid of the mat or the pier itself can be battered to the tower leg slope.

The stub angle is embedded in the top of the pier so that the upper exposed section can be spliced directly to the main tower leg and diagonals. The stub angle should be of adequate size to resist the

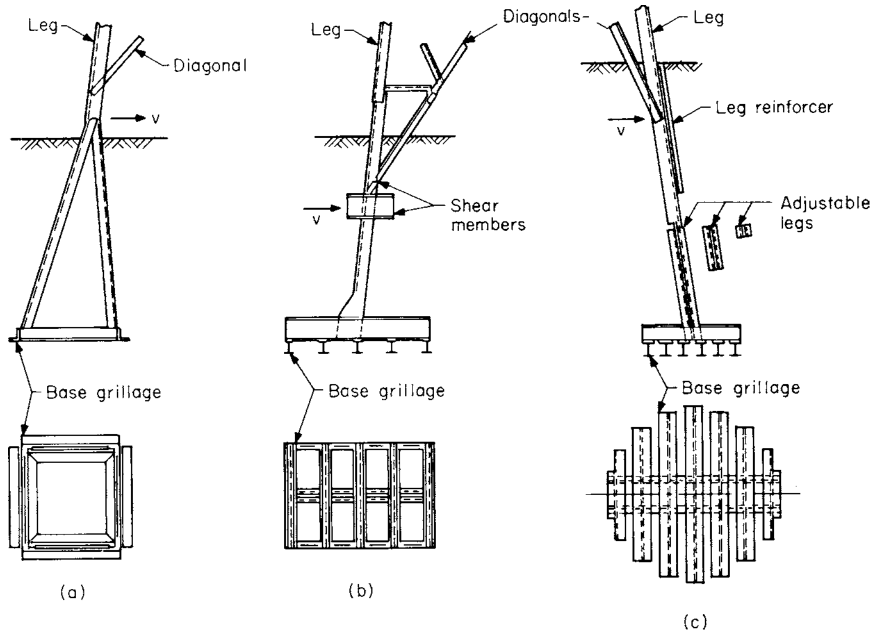


FIGURE 14-64 Typical steel grillage foundations.

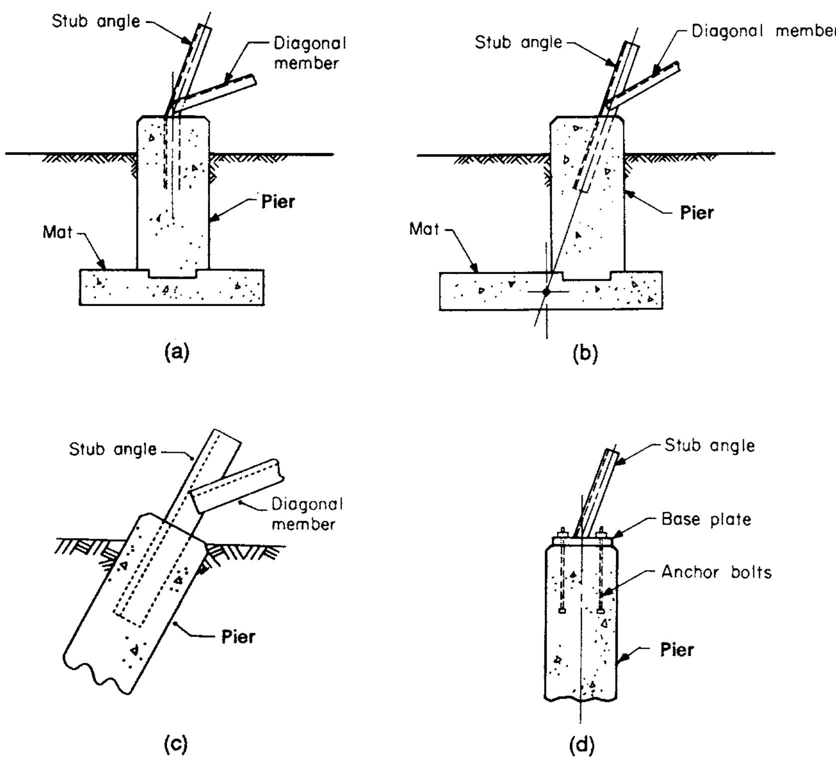


FIGURE 14-65 Typical concrete spread foundations (reinforcing not shown).

axial loads transmitted from the main leg and diagonals plus any secondary bending moment from the horizontal shear, if applicable. The stub angle must be embedded in the concrete to a sufficient length to transmit the load to the concrete. Bolted clip angles or welded stud shear connectors may be added to the lower end of the stub to reduce its length. Anchor bolts can also be used in lieu of the direct embedment stub angle as shown in Fig. 14-65*d*.

Rock Foundations. Many areas of the United States have bedrock either exposed at the ground surface or covered with a thin mantle of soil. Relatively simple, economical, and efficient rock foundations may be installed where this type of terrain is encountered. A rock foundation can be designed to resist both uplift and compression loads plus horizontal shear and, in some structure applications, bending moments. Where suitable bedrock is encountered at the surface or close to the surface, a rock foundation, as shown in Fig. 14-66, can be installed. Bolted clip angles or welded stud shear connectors may be added to the lower end of the stud to reduce its length.

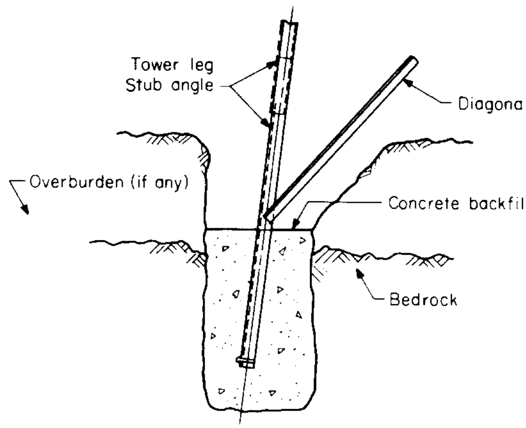


FIGURE 14-66 Rock foundations.

Drilled Shafts. The drilled concrete shaft is the most common type of foundation presently being used to support lattice towers, framed structures, and single shafts. Drilled concrete shafts are constructed by power augering a circular excavation, placing the reinforcing steel and pouring concrete to form a shaft foundation. Lattice towers are attached by embedment of a stub angle or through the use of baseplates and anchor bolts. Framed structures and single shafts are attached through the use of baseplates and anchor bolts.

Drilled shafts can be constructed in a wide variety of soil and rock types. However, the construction of drilled concrete shafts may encounter problems under certain soil conditions. For example, granular soils may collapse into the excavation before concrete can be poured. In soft, cohesive soils, squeezing or shear failure of the soil can occur, producing a reduced diameter or the excavation may become completely obstructed before the concrete is placed. This soil movement in the excavation can result in ground-surface settlement. Thus, casing and/or drilling mud may be required in granular and soft cohesive soils to maintain an open excavation.

Direct Embedment. Direct embedment refers to wood, steel, or concrete pole foundations (both single-shaft and H-framed structures) constructed by power augering a circular excavation in the ground, inserting the pole directly into the excavation, and backfilling the void between the pole and the sides of the excavation. Thus, the pole acts as its own foundation by transferring loads to the in situ soil via the backfill. This technique has been traditionally used for wood-pole foundations and has recently been employed for metal- and concrete-pole foundations.

The quality of backfill, method of placement, and degree of compaction strongly influence the stiffness and strength of a direct embedment foundation. Corrosion of an embedded metal pole is also an important consideration. It should be noted that the presence of granular or soft, cohesive soils may cause the same construction problems for direct embedment foundations as for drilled concrete-shaft foundations.

Subsurface Investigations. The technical requirement of assuring a safe and cost-effective foundation design for transmission structures requires a thorough knowledge of the subsurface conditions along the right-of-way (ROW). The intent of this section is to provide a guide for performing an adequate subsurface investigation for the design of transmission-line structure foundations.

When designing a foundation, the engineer should be concerned with the following factors: (1) the ultimate load-bearing capacity of the subsurface material, and (2) the allowable displacements of the foundation. In the case of grillages and directly embedded poles, the quality of available backfill materials is also a concern. Hence, the objectives of a subsurface investigation are to determine the stratigraphy, physical characteristics, and engineering properties (particularly the strength and deformation characteristics) of the soil or rock underlying a given site.

To determine the most cost-effective foundation, it is necessary to consider the engineering and physical properties of the subsurface materials; construction costs; the construction aspects of a particular foundation type and how they are influenced by such factors as groundwater elevation, safety requirements, contractor capability and experience, and environmental constraints.

The scope of a subsurface investigation will vary depending on foundation loads, type of structure, and probable foundation types, types of subsurface materials, and previous knowledge of subsurface conditions along the line route. It is necessary to use engineering judgment when considering the scope of the subsurface investigation. A detailed outline for developing a cost-effective investigation is given in an EPRI report.[99]

Design of Spread Foundations. Spread foundations are used to support lattice towers. The design of spread foundations for transmission towers must consider both the direction (uplift or compression) and orientation (inclination and eccentricity) of the applied loads. The foundation must be designed to prevent excessive displacement or shear (bearing capacity or uplift) failure of the support soil. A detailed presentation of estimating the uplift and compression capacity of spread foundations is given in an EPRI report.[99]

Design of Drilled Shaft Foundations. Drilled shaft foundations are used to support lattice-tower, framed, and single-shaft structures. This type of foundation supports vertical compression loads through a combination of side shear and end bearing and supports vertical uplift loads by side shear. Lateral loads and overturning moments are supported by lateral resistance of the soil and/or rock in which the shaft is embedded plus the vertical shearing resistance on the perimeter of the shaft, and the horizontal shear on the base and the base moment.

Compression and Uplift Capacity. Methods for computing the compression and uplift capacity of drilled shaft foundations are given in an EPRI report.[100]

Lateral Load Capacity. The response of a drilled shaft to lateral loads is the result of complex interactions between the shaft and the soil and/or rock in which it is embedded. A common method of modeling this interaction is called the subgrade modulus approach. Reference 101 provides a detailed explanation of a method for determining the lateral capacity of drilled shafts. A computer program, MFAD (Moment Foundation Analysis and Design), originally called PADLL, which was developed as part of the EPRI research project eliminates the simplifying assumptions associated with prior models.

Direct Embedment. The response of direct embedment foundations in compression, uplift, and lateral loads is similar to that of drilled concrete shafts. Most of the analytical techniques used in drilled shaft design are relevant to direct embedment design. The principal differences between direct embedment foundations and drilled concrete shaft foundations are: (1) the backfill which intervenes between the pole and the in situ soil, and (2) the stiffness of the embedded structure shaft relative to that of a drilled concrete shaft. Drilled shafts transfer loads directly to the in situ soil. A detailed presentation for the analysis and design of directly embedded poles is given in an EPRI report.[102] Results from 12 full-scale tests on single-pole direct embedment foundations are given in a related report.[103]

Construction Considerations. Factors affecting, and problems associated with, the construction of drilled shaft foundations can be divided into two general areas: (1) geotechnical factors influencing construction, and (2) construction-related problems. A common geotechnical occurrence is the erosion (sloughing, caving in) of loose granular soil layers, mainly below the groundwater level. Another geotechnical factor influencing the overall capacity of the foundation is the release of stresses due to excavation, especially when the hole is left open for a long period of time (say, one or more days, depending on soil conditions). Some construction-related problems are associated with the use of drilling mud (slurry method), which tends to leave an undesirable film (up to various inches of thickness) of soft material adhered to the walls of the hole. Another frequently occurring problem is the perturbance and remolding generated by the use of casing. Also, problems arise when special geometry is requested by the designer for the drilled shaft (under-ream, shear rings, etc.). In all of these cases, in general, good communication is required between field personnel and designers to permit early detection of these conditions.

Design of Anchors and Anchor Foundations. An anchor is a device which will provide resistance to an upward (tensile) force transferred to the anchor by a guy wire or structure leg member. An anchor may be a steel plate, wooden log, or concrete slab buried in the ground, a deformed bar or a steel cable grouted into a hole drilled into either soil or rock, or one of several manufactured anchors which are either drilled or rotated into the ground. Anchorage may also be provided by vertical or battered drilled shafts or piles. Typical types of anchors are shown in Fig. 14-67.

Anchors may be classified as either deadman or prestressed. Deadman anchors are defined as those anchors which are not loaded until the structure is loaded. Prestressed anchors are loaded to specified load levels during installation of the anchor. An advantage of a prestressed anchor is that most of the initial strains of the anchorage system have been removed before the structural load is applied. Therefore, the full capacity of the anchor can be attained at very small deformation (movements in soil of less than $^1/_4$ in are typical). Another advantage of prestressed anchors is that they are proof-loaded to their design load at the time of installation. Disadvantages of prestressed anchors are

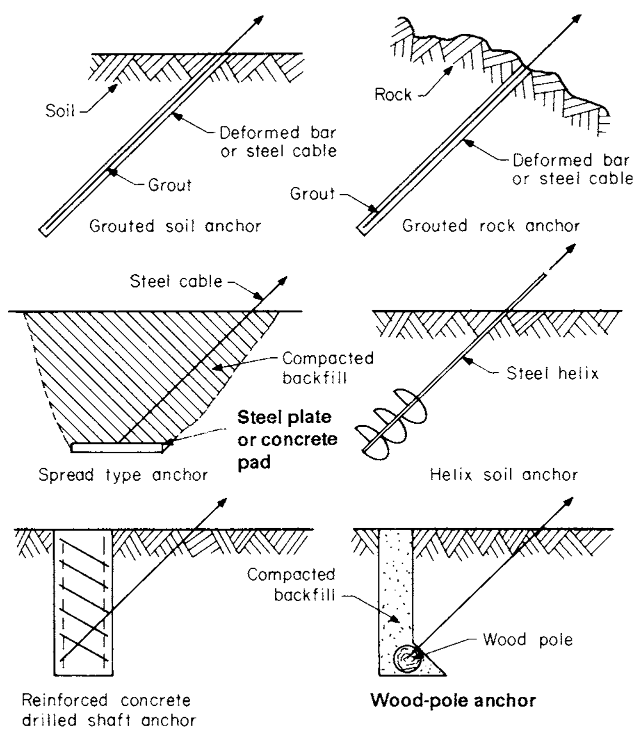


FIGURE 14-67 Typical anchors.

that they are generally more expensive than deadman anchors and they should not be used in compressible soils.

Another advantage of prestressed anchors is that shallow anchors may obtain additional strength by the increased effective stress created by the influence of the bearing plate on the soils adjacent to the anchors.

Deadman anchors may include any of the systems shown in Fig. 14-67. Initial strains in deadman anchors may be reduced by as much as 50% by prestressing them to their design load at the time of installation.

Anchor Application. Anchors are used to permanently support guyed structures, as well as to temporarily support other structure types during erection and stringing. The legs of lattice towers can be anchored directly by rock anchors or helix-type anchors. The uplift capacity of spread foundations may be increased through the use of anchors as shown in Fig. 14-68. Guys and anchors are also extensively used to terminate wire loads on wood structures and to increase wood structure capacity for high transverse loading. At intermediate structure locations, guys and anchors may be utilized to provide additional longitudinal strength. Anchors can be used to increase the load capacity of existing foundations.

Design. The design of an anchor depends on a knowledge of the peak and residual shear strength properties of the soil or rock in which it is embedded. In rock, it is also important to know the degree and depth of any weathering which may have occurred, together with the orientation and spacing of joints and foliation. In addition, an understanding of the load characteristics and the structure deflection tolerance combined with the guy cable elongation is important in selecting and designing the type of anchor. Anchor pullout tests are often conducted to confirm design assumptions where prior experience is lacking.

References 99, 104, and 105 provide detailed information on the design of anchors.

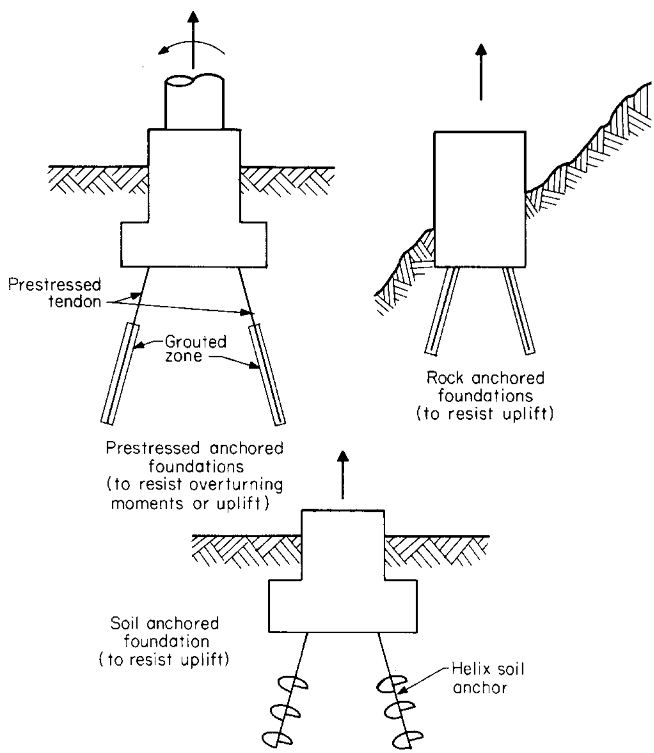


FIGURE 14-68 Typical anchored spread foundations.

14.1.15 Overhead Line Uprating and Upgrading

Performing a transmission line uprating can be very attractive in terms of getting smaller costs and shorter leadtimes when compared to building a new line. Besides that, it can postpone the need of new lines, reduce congestion costs and avoid unnecessary load shedding during contingencies. However, before starting the uprating task, it is important to evaluate some feasibility issues, as well as to choose the most appropriate type of uprating (Thermal Uprating or Voltage Uprating) for a specific transmission line. Each type of uprating requires different kinds of previous analysis and shall be done by its own methods. Sometimes a transmission line uprating also requires a line upgrading. Transmission line upgrading is related to physical modifications to the line.[106,107]

Uprating Feasibility Issues

Technical Feasibility. For this kind of analysis it is important to consider at least the following points:

- *System load requirements*. It is important to evaluate for how long the uprated/upgraded line will satisfy the load requirements.
- *Assessment of current conditions and life expectancy of transmission line materials.* It is important to make this kind of evaluation for the main transmission line components, such as towers, foundations, conductors, insulators, and hardware.
- *Potential margins for uprating/upgrading.* It is important to check electrical clearances, mechanical strengths, ROW width, as well as the possibility of compliance with the requirements of safety codes (e.g., NESC), regulatory bodies and government agencies (e.g., navigable streams, public lands, air lanes).
- *Utility considerations.* Sometimes electric utilities are not authorized to take the transmission line out of service to perform the necessary uprate/upgrade services. In these cases it is important to check if the mentioned services can be done with the line in service.

Economical/Financial Feasibility. For this kind of analysis it is important to consider at least the following points:

- Uprating/upgrading costs vs. new line costs. It is important to remember that technical analysis of old lines usually requires data gathering and this can be very expensive and time consuming. Besides that, it is necessary to estimate what will be the need of the uprated line in terms of additional ROW. Other costs that can be relevant are related to construction (material and labor), maintenance and operation of the uprated line. Environmental costs are usually higher for new lines.
- Uprating/upgrading costs vs. uprating/upgrading benefits.

Environmental Feasibility. For this kind of analysis it is important to consider at least the following points:

- *Environmental considerations.* Usually not so critical when compared to new lines. However, it may be necessary to deal with historical societies, environmental groups, concerned neighbors, and so forth.
- Right-of-way easements. If significant changes will be made to the original line, it is necessary to check the validity of the previous ROW terms of use. It can be difficult to get licensing for the modified line. It is also important to check the existence of ROW encroachments and line crossings that would be unacceptable by the uprated line.

Thermal Uprating

Effectiveness. This kind of uprating can be a good option when the line loading is limited by thermal constraints and the line has margin in terms of maximum allowable conductor temperature.

Previous Analysis to Perform. Before proceeding with a transmission line thermal uprating it is necessary to analyze maximum allowable conductor temperature, conductor-to-ground clearance, magnetic fields, ROW issues, and sometimes structural strengths.

Usual Techniques. Some of the techniques used to perform transmission line thermal uprating are described as follows:

- Performing dynamic thermal rating monitoring
- Raising the thermal limit imposed to the line by some inexpensive substation equipment (e.g., disconnect switches)
- Raising the thermal limit by making similar the thermal limits of all line sections
- Keeping appropriate conductor-to-ground clearances while increasing the maximum allowable conductor temperature
- Bundling the original line conductor with another one, or replacing the line conductor by a more conductive one, to increase the line current-carrying capacity

Some of these techniques have large structural impacts. Performing a transmission line thermal uprating can bring some impact to the substation equipment.

Voltage Uprating. This kind of transmission line uprating can result in a much higher rating increase than thermal uprating. Besides that, transferring the same amount of power in a higher voltage level reduces the line current, and consequently, line losses and voltage drops. However, voltage uprating is typically more expensive than thermal uprating due to the need of also uprating the voltage class of the terminal substations equipment.

Effectiveness. This kind of uprating can be a good option when: the line loading is limited by voltage drop or stability considerations; the line has margins in terms of electrical clearances; the uprating can be done with minimal line modifications or it will be applied to several circuits simultaneously, or the line design criteria can be relaxed.

Previous Analysis to Perform. Before proceeding with a transmission line voltage uprating it is necessary to analyze tower clearances, conductor-to-ground clearance, corona performance, electric fields, ROW issues, and sometimes structural strengths.

Some Usual Voltage Uprating Techniques. Some of the techniques used to perform transmission line voltage uprating are described as follows:

- Addition of insulator units to the transmission line insulator strings
- Replacement of standard insulator units by polymeric or antifog units
- Application of strut insulators (or V strings) to prevent swinging of suspension strings
- Keeping appropriate conductor-to-ground clearances while increasing the transmission line operating voltage
  - Retensioning the existing conductors
  - Performing sag adjustments (cutting out conductor lengths, sliding conductor clamps)
  - Increasing the conductor height at the attachment support (converting suspension strings to pseudo dead-end strings)
  - Increasing the attachment support height
  - Raising towers
  - Moving towers
  - Inserting additional towers
  - Performing terrain contouring (rural areas)
- Bundling the original line conductor with another one, or replacing the line conductor by a bigger one, to assure a good corona performance
- Performing Line Compaction
- Converting a 3-phase double-circuit line to a 6-phase single-circuit line
- Converting a low voltage double-circuit line to a high-voltage single-circuit line
- Converting HVAC lines to HVDC lines

Some of these techniques have large structural impacts.

Voltage Upgrading of Existing Lines. Figure 14-69 illustrates the fact that as transmission voltages increased over the years experience and research has allowed smaller conductor spacings in per unit of the spacing required to withstand power frequency voltage flashover. An observation from Fig. 14-69 is that many lines designed years ago may have clearances sufficient for operation at a higher voltage than the initial operating voltage of the line. Thus, possibilities exist for upgrading the voltage of some older existing transmission lines.

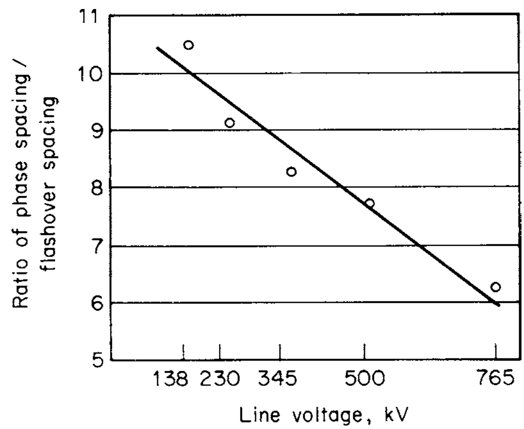


FIGURE 14-69 Phase-spacing ratio versus line voltage.

A voltage upgrading study involves reengineering the line using the best available data. Upgrading differs from new line design because operating experience with the existing line provides a source of data not available to the designer of a new line. It also differs from new line design in that certain reduced design margins may be accepted in order to avoid a complete replacement of structures and conductors.

Voltage upgrading of transmission line ranges from increasing the voltage with minimal modifications to virtual reconstruction of the line. The greater the extent of the modifications, the more mechanical and structural issues become significant. An important part of an upgrading study is careful verification of the existing condition of the line, including conductors, insulators, geometry, structures, and foundations.

Insulation concerns are paramount in a voltage upgrading analysis but trade-offs are frequently necessary between competing requirements. For example, it may not be possible to add insulators of the same type as the original while maintaining the same leakage distance per kilowatt and without adversely impinging on air gap clearances. In such cases several options may be considered, including relaxation of the initial design criteria.

Insulation for power frequency voltage is concerned with contamination performance of insulators and air gaps between energized and grounded line components. An advantage of an upgrading study is the availability of operating experience of the existing line. It is important to carefully review the operating performance of the line at the existing voltage, because problems that may have been previously ignored may become significant with the increased electrical stress at the higher voltage.

Insulator leakage distance required is a function of the degree of insulator contamination at the location of the line. The degree of contamination maybe well known, or if necessary can be determined by laboratory tests. An insulator string can be carefully removed from the existing line so as not to disturb the surface dirt and sent to a laboratory for testing in a fog chamber, or alternatively a sample of the contamination can be taken from insulators on the line and the equivalent salt deposit density determined.

Once the degree of contamination is understood, several options can be considered if there is insufficient space to add conventional insulator units to the strings:

- *Use less leakage distance per kilowatt*. It may be determined that adequate performance can be obtained with less leakage distance than would be customarily be applied to a new line design.
- *Different insulator types*. Additional leakage distance may be possible by application of high leakage or fog-type insulators without excessively lengthening the insulator string length. The hydrophobic properties of polymer insulators may be considered for reducing the leakage distance. Porcelain insulators with semiconducting glaze provide improved contamination performance and may be an option.

Adequate air gap clearances must be maintained under maximum wind swing conditions. Because power frequency voltage is always present, it is necessary to consider insulator swing for the maximum predicted wind speed to maintain adequate clearance for power frequency voltage withstand. It is also necessary to maintain adequate clearances to the edge of the ROW to meet code requirements.

Switching surges have typically not been an issue for the lower transmission voltages. However, in a voltage upgrading, it is necessary to utilize the available space to the maximum degree possible and switching surges can become the dominant effect in limiting voltages. Therefore, a careful switching surge analysis is an important part of a voltage upgrading study. Some transmission lines have been successfully upgraded with proper control of switching surge overvoltages. Control methods may be necessary even at 115 kV and include:

- Use of resistor preinsertion in the circuit breakers
- Application of surge arresters
- Synchronized pole closing circuit breakers

Code and maintenance issues are part of a voltage upgrading study. Control of switching surge overvoltages may allow operation of the line at the new voltage with the same ground clearance that existed previously. The old line may have been designed in compliance with older code editions that are based on a linear relationship between clearance and voltage. Control of switching surge overvoltages may allow use of the alternate clearance calculation and reduced clearances compared to the linear calculation. Likewise, switching surge overvoltages are part of the safety considerations for live line working.

If it is necessary to increase conductor ground clearance to maintain code clearances, retensioning of the conductors may be a possible way of meeting the code requirements. Retensioning conductors increases mechanical forces on dead end and heavy angle structures and necessitates analysis of structure loading. Code ground clearances may also relate to the allowable length of insulator strings. Structures may be modified to increase conductor height. Once structure modifications are considered, structure and foundation loadings must be carefully studied.

Lightning tripout performance is seldom a limiting consideration for a voltage upgrading study. Lengthening the insulator string increases the insulator impulse CFO. Increased voltage adds to the lightning impulse voltage half the time and subtracts half the time depending on the instant of the lightning strike in relation to the power frequency voltage sine wave. The degree of change of line geometry depends on the amount of structural modifications necessary. A lightning performance calculation program can be used to compare the performance before and after the upgrading. Comparison of the calculated performance of the existing line with the line operating record calibrates the program to the line and allows estimation of the change in performance to be expected with the upgrading. Normally the change will be slight. If it is desired to improve the lightning performance with the upgrading, methods such as reducing the structure footing resistance are available.

Corona on energized conductors, hardware, and insulators is a fundamental limitation on the maximum possible voltage on any transmission line. As the voltage is increased, the electric field at the surface of energized components increases. At some point the electric field becomes sufficient to ionize the air and produce excessive corona. Corona manifests itself as electromagnetic interference (EMI—radio noise and television interference), audible noise, and in more extreme cases visible corona and corona power loss. An analysis of corona effects is an essential part of a voltage upgrading study.

Audible noise traditionally has only been a concern for the higher transmission voltages, for example, 500 kV and above. Because corona effects are often the limiting factor for a voltage upgrade, the conductor surface electric field on an upgraded line can be at levels typical of EHV lines, even for 115 kV. Thus, audible noise can become an issue for upgraded lines at voltages where audible noise has traditionally not been a consideration. In any event, it may be necessary to demonstrate adequate audible noise performance to obtain authorization for the upgrade.

Conductor corona caused EMI is the quantity predicted by radio and television noise programs and is the value traditionally used for evaluation of new transmission line designs. The assumption is made that if noise from conductor corona is at an acceptable level, noise from other sources will be insignificant. If EMI limits the voltage below the desired new voltage, mitigation options are reconductoring or adding a second conductor per phase. The second conductor per phase may be added in either a horizontal or vertical configuration. The vertical configuration may be preferred if it impinges less on the air gap clearances to the structure under wind swinging than the horizontal configuration. Reconductoring or adding a second conductor per phase requires a mechanical and structural analysis to ensure the conductors can be properly supported.

While conductor corona EMI is also an essential consideration for voltage upgrading, additional noise sources such as corona on hardware and insulators may become important. "Corona-free" hardware has been used for EHV lines where the electric field on the hardware is greater, while standard hardware has been used for lower voltage lines. The difference for bolted conductor shoes is that standard hardware has the U-bolts arranged so the nuts are on the underside of the shoe, while for corona-free shoes the U-bolts are inverted so the nuts are inside the shoe where they are electrically shielded. Because the electric field on the hardware of an upgraded line may be at levels typically associated with EHV lines, it becomes necessary to change the hardware to corona-free hardware to avoid excessive radio noise.

Under some conditions it may be desired to test hardware or insulators to ensure adequate corona performance. If laboratory tests are performed, care must be taken to properly reproduce the expected electric field at the surface of the hardware or insulators. Traditional laboratory tests have used a voltage above the desired operating voltage and measurement of radio influence voltage (RIV). These tests have been adequate for traditionally designed lines, but may give misleading results for compact or upgraded lines where the surface electric fields are greater than traditional values for the voltage level. Therefore it is necessary to properly characterize the electric field on the device under test.

Experience has indicated that when a transmission line is properly designed for corona-caused EMI, more than 90% to 95% of radio noise complaints and virtually all television interference complaints are a result of spark sources. These spark sources are located and repaired by normal maintenance techniques. Such spark sources include such things as loose bolts or down lead staples on wood poles or films of corrosion between caps and pins of adjacent units in slack strings of suspension insulators. Noise complaint experience for the line considered for upgrading and the physical condition of the line are valuable indicators of the possible presence of spark noise sources. Increasing the line voltage will increase the possibility of spark noise sources and should be considered as part of a voltage upgrading study.

Upgrading voltage might be considered for a line remote from population where audible and EMI noise are deemed of secondary importance. In such a case the possibility of upgrading may be limited by visible corona and corona loss. At some point, as the voltage is increased the power loss in foul weather becomes excessive and increase of voltage becomes impractical.

Voltage upgrading studies are frequently performed in two parts: a preliminary analysis to determine if voltage upgrading of the desired line is feasible, followed by a detailed analysis if the feasibility analysis proves promising. Typical or approximated data may be used to assess feasibility. The scope of the detailed analysis is set by the feasibility study. Not all voltage upgrading studies include all of the same elements. Laboratory tests may be included, in some cases, even construction of a prototype line section for testing. In other cases, some aspects of the detailed study may be of smaller importance because of the degree of margin in the existing design.

Compact Line Design. The most fruitful areas for line compaction are 69- to 230-kV lines, because these designs are generally older, with more generous clearances. Figure 14-70 illustrates a comparison of compact and conventional 230-kV structures. Some compaction is possible, however, at higher transmission voltages by redesign, as shown in Fig. 14-71.

Compact line design techniques may be applied for more efficient use of ROW for new construction, voltage uprating of existing lower-voltage lines, and possible economies. Line compaction is also a tool for reducing electric and magnetic fields at ground level.

A compact transmission-line design analysis consists of the following topics:

Conductor Surface Electric Field and Corona. A given conductor in air has a maximum surface electric field for acceptable levels of corona. EHV lines are typically designed for conductor surface electric fields near this maximum. Compaction of lower-voltage lines raises the conductor surface electric field to EHV levels. Thus conductor and phase spacing/configuration must take corona

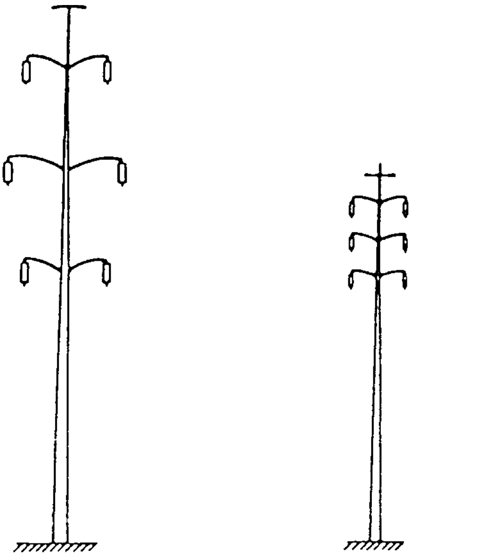


FIGURE 14-70 Comparison of conventional and compact 230-kV lines.

phenomena into account: corona loss, radio noise, and audible noise. Audible noise is not a normal consideration for lower transmission voltages, but should be analyzed for compact lines because of the field levels involved. Because the conductor surface electric field of a compact line is at EHV levels, it is necessary to use EHV-class hardware to limit hardware corona. Laboratory test procedures must be specified in terms of electric field on the conductor surface rather than system voltage.

Insulation Performance. Power frequency voltage (including insulator contamination), switching surge, and lightning performance must be evaluated for the proposed design. Switching surge control can be provided either by surge arresters or resistor preinsertion in the circuit breakers.

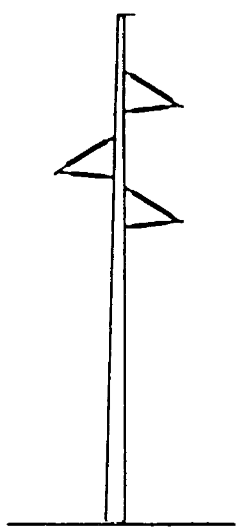


FIGURE 14-71 A 500-kV compact line.

Conductor Motion. The one consideration in design of a compact high-voltage transmission line which differs from EHV practice is mechanical motion of the conductors under wind, ice, and through-fault currents. It is imperative that conductor motion be limited to maintain adequate clearances for power frequency voltage. Differential wind motion, conductor motion under ice release, and ice galloping are extensions of existing practice. A new consideration is motion which results from *through-fault current* (fault current which flows on a line due to a fault somewhere else on the system). Especially with horizontal conductor configurations, magnetic forces from through faults can develop enough conductor motion to cause a line trip.

Because of the need to restrict conductor motion, compact lines are customarily designed with post insulators to restrict conductor motion at the structures. If necessary, inspan insulating spacers can be applied to restrict the conductor motion. Maximum compaction (minimum conductor spacing) is achieved by designs which eliminate grounded supporting structure members in the space between conductors. The spacing reduction is a result of the need for clearances for phase-to-phase voltage between conductors rather than twice the phase-to-ground clearance. Portal structure designs support the conductors from outside the conductor array and are useful for line compaction.

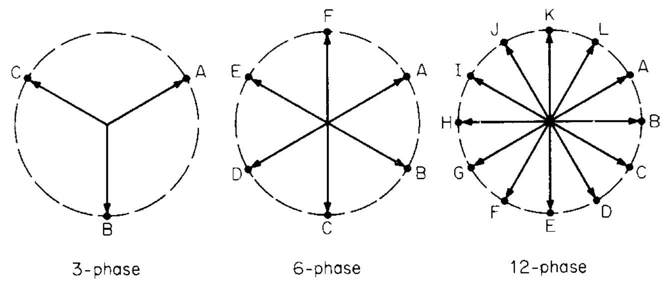


FIGURE 14-72 Illustration of high-phase-order phasor diagrams.

A limitation of compact line design may be a restriction on live-line maintenance because of the reduced clearances. If live-line maintenance is considered in the design phase, it may be possible to develop procedures and to modify the design to allow live-line work while meeting applicable safety codes.[108]

High-Phase-Order Transmission.[109–112]. This is an extension of the principles of line compaction to more than three phases. A general principle can be developed for maximum power transfer in a given amount of space with a given amount of conductor material. The result of this development is that the conductors should be spread symmetrically around the periphery, and energized with the highest possible voltages and phase angles corresponding to the space angle between conductors. The conductor and phase configuration is illustrated for 6- and 12-phase in Fig. 14-72.[111] High-phase-order lines can be constructed smaller than 3-phase lines for equivalent capacity.

An alternative way of looking at high phase order is that conversion of a double-circuit 3-phase line to 6-phase at the same phase-to-ground voltage reduces the conductor surface electric field. This reduced conductor surface electric field can be used either to bring the conductors closer together (compaction) or to increase the phase-to-ground voltage. Thermal loading increases linearly with voltage, and surge impedance loading increases with the square of the voltage. As a result, high phase order holds the potential for voltage uprating of existing double-circuit 3-phase lines by conversion to 6-phase at a higher phase-to-ground voltage.

The benefits and practicality of high phase order have been proved by 3 years' successful operation of a demonstration 6-phase line near Binghamton, New York. An existing double-circuit 115-kV line was converted to 93-kV 6-phase, for a 40% increase in phase-to-ground voltage. A section was converted to compact design, and relay protection was developed and tested. Substation layout proved practical, and economic results are favorable for appropriate applications.[114–118]

REFERENCES

1. Barnes, H. C., and Thoren, B., The AEP-ASEA UHV Test Station and Line, IEEE Conf., paper C73-319-1, presented at IEEE Summer Power Meeting, 1973.
2. Development of Ultra-high Voltage Transmission, Bonneville Power Administration, Portland, Ore., July 1974.
3. *Transmission Line Reference Book, 345 kV and Above*, 2nd ed., Electric Power Research Institute, Palo Alto, Calif., 1982.
4. Douglass, D. A., Economic Measures of Bare Overhead Conductor Characteristics, *IEEE Trans. Power Delivery*, April 1988, vol. 3, no. 2.
5. Douglass, D. A., and Kennon, R. E., EHV Transmission Line Design Opportunities for Cost Reduction, *IEEE Trans. Power Delivery*, April 1990, vol. 5, no. 2.
6. Bates, J., and White, H. B., Micro-Based Program Refines Transmission Design, *Electrical World Magazine*, April 1987.

7. Clarke, E., *Circuit Analysis of A-C Power Systems*, Wiley, New York, 1943.
8. Keast, D. N., Assessing the Impact of Audible Noise from AC Transmission Lines: A Proposed Method, *IEEE Trans. Power Appar. Syst.*, May/June 1980, vol. PAS-99, no. 3, p. 1021.
9. Clayton, R. E., and Stewart, J. R., Transmission Line Electromagnetic Compatibility, 1975 IEEE Electromagnetic Compatibility Symposium Record, IEEE publication 75CH1002-5EMC.
10. A Comparison of Methods for Calculating Audible Noise of High Voltage Transmission Lines, IEEE Task Force Report, *IEEE Trans. Power Appar. Syst.*, Oct. 1982, vol. PAS-101, no. 10, p. 4290.
11. Perry, D. E., An Analysis of Transmission Line Audible Noise Based upon Field and Three-Phase Test Line Measurements, *IEEE Trans. Power Appar. Syst.*, May/June 1972, vol. PAS-91, p. 857.
12. Measurement of Audible Noise from Transmission Lines, IEEE Task Force Report, *IEEE Trans. Power Appar. Syst.*, March, 1981, vol. PAS-100, no. 3, p. 1442.
13. Loftness, M., *AC Power Interference Manual*, Percival Publishing, Tumwater, Wash.
14. Chartier, V. L., et al., Investigation of Corona and Field Effects of AC/DC Hybrid Transmission Lines, *IEEE Trans. Power Appar. Syst.*, Jan. 1981, vol. PAS-100, no. 1, p. 72.
15. Review of Technical Considerations on Limits to Interference from Power Lines and Stations, IEEE Committee Report, *IEEE Trans. Power Appar. Syst.*, Jan./Feb. 1980, vol. PAS-99, no. 1, p. 365.
16. *Tolerable Limits and Methods of Measurement of Electromagnetic Interference from Alternating Current High Voltage Power Systems*, CSA Standard C108.3.1-1975, Canadian Standards Association.
17. *Human Response to Interference with TV Picture Quality*, Report EL-1587, Project 68-4, Electric Power Research Institute, Palo Alto, Calif., 1980.
18. Scherer, H. N., Jr., Ware, B. J., and Shih, C. H., Gaseous Effluents Due to EHV Transmission Line Corona, paper T 72 550-2, *IEEE Trans. Power Appar. Syst.*, May/June, 1973, vol. PAS-92, no. 3.
19. Frydman, M., Levy, A., and Miller, S. E., Oxidant Measurements in the Vicinity of Energized 765 kV Lines, IEEE paper T 72 551-0, *IEEE Trans. Power Appar. Syst.*, May/June, 1973, vol. PAS-92, no. 3.
20. Fern, W. J., and Brabets, R. I., Field Investigation of Ozone Adjacent to High Voltage Transmission Lines, IEEE paper T 74 057-6, *IEEE Trans. Power Appar. Syst.*, Sept./Oct. 1974, vol. PAS-93, no. 5.
21. Frydman, M., and Shih, C. H., Effects of the Environment on Oxidants Production in AC Corona, IEEE paper T 73 407-4, *IEEE Trans. Power Appar. Syst.*, Jan./Feb., 1974, vol. PAS-93, no. 1.
22. Roach, J. F., Chartier, V. L., and Dietrich, F. M., Experimental Oxidant Production Rates for EHV Transmission Lines and Theoretical Estimates of Ozone Concentrations Near Operating Lines, IEEE paper T 73 414-0, *IEEE Trans. Power Appar. Syst.*, March/April, 1974, vol. PAS-93, no. 2.
23. Abel, W. A., Comparison of Ozone Instrumentation, IEEE paper A 78 166-7, abstract in *IEEE Trans. Power Appar. Syst.*, July/Aug. 1978, vol. PAS-97, no. 4, p. 1009.
24. Roach, J. F., et al., Ozone Concentration Measurements on the C-Line at the Apple Grove 750 kV Project and Theoretical Estimates of Ozone Concentrations Near 765 kV Lines of Normal Design, *IEEE Trans. Power Appar. Syst.*, July/Aug. 1978, vol. PAS-97, no. 4, p. 1392.
25. Sebo, S. A., et al., Examination of Ozone Emanating from EHV Transmission Line Corona Discharges, *IEEE Trans. Power Appar. Syst.*, March/April 1976, vol. PAS-95, no. 2.
26. *The Electrostatic and Electromagnetic Effects of AC Transmission Lines*, IEEE Tutorial 79 EH 0145-3 PWR, 1979.
27. Electric and Magnetic Field Coupling from High Voltage AC Power Transmission Lines-Classification of Short-Term Effects on People, IEEE Committee Paper, *IEEE Trans. Power Appar. Syst.*, Nov./Dec. 1978, vol. PAS-97, no. 6, p. 2243.
28. Chiu, M. C., Fuel Ignition by High Voltage Capacitive Discharges, report JHU PPSET-18, March 1983, John Hopkins University Applied Physics Laboratory, Laurel, Md.
29. Hamaam, M. S., and Baishiki, R. S., A Range of Body Impedance Values for Low Voltage, Low Source Impedance Systems of 60 Hz; *IEEE Trans. Power Appar. Syst.*, May 1983, vol. PAS-102, no. 5, p. 1097.
30. Jaffa, K. C., Magnetic Field Induction from Overhead Transmission and Distribution Power Lines on Parallel Fences, *IEEE Trans. Power Appar. Syst.,* April 1981, vol. PAS-100, no. 4, p. 1624.
31. Dabkowski, J., The Calculation of Magnetic Coupling from Overhead Transmission Lines, *IEEE Trans. Power Appar. Syst.*, Aug. 1981, vol. PAS-100, no. 8, p. 3850.
32. Taylor, R. J., Hazard Analysis for Magnetic Induction from Electric Transmission Lines, report JHU PPSE T-23, March 1982, John Hopkins University Applied Physics Laboratory, Laurel, Md.

33. *Procedures for Measurement of Power Frequency Electric and Magnetic Fields from AC Power Lines.* IEEE Standard 644, IEEE, New York, 1994.
34. Kaminski, J., Jr., Long Time Mechanical and Electrical Strength in Suspended Insulators, *Trans. AIEE*, Aug. 1963, p. 446.
35. Nicholas, F. S., and Vose, F. C., A Polymer Insulator for High Voltage Transmission Lines, *Elec. Eng.,* vol. 82, 1963.
36. Abilgaard, E. H., Bauer, E. A., et al., Composite Longrod Insulators and Their Influence on the Design of Overhead Lines, CIGRE paper 22-03, 1976.
37. Development of Polymer Bonded Silica, (Polysil) for Electrical Applications, EPRI Report EL488, May, 1977.
38. The Metapol Insulator: Dulmison (Australia) Inc. Catalog.
39. Karady, G., and Lamontagne, G., Electrical and Contamination Performance of Synthetic Insulators for 735 kV Transmission Lines, IEEE paper A76 502-5, presented at IEEE PES Summer Meeting, Portland, Ore., July 18–23, 1976.
40. Broschat, M., Transmission Line Uprating 115 kV to 230 kV, Report on Operating Performance, *IEEE Trans.*, March/April 1972, pp. 545–548.
41. Update Line to 345 kV on same ROW; *Electr. World*, Nov. 15, 1973, pp. 66–67.
42. *EHV Transmission Line Reference Book*, EEI, 1968.
43. IEEE Working Group on Insulator Contamination, Application Guide for Insulators in a Contaminated Environment, *IEEE Trans. Power Appar. Syst.*, Sept./Oct. 1979, pp. 1676–1690.
44. *Transmission Line Reference Book, 115-138 kV Compact Line Design*, EPRI Publication, 1978.
45. Moran, J. H., The Effect of Cold Switch-on on Semi-conducting Glazed Insulators, IEEE paper C74 071-7, presented at IEEE Power Meeting, New York, Jan., 1974.
46. Moran, J. H., and Powell, D. G., A Possible Solution to the Insulator Contamination Problem, IEEE paper 71CP41, presented at IEEE Power Meeting, New York, Jan., 1971.
47. Falter, S. L., and Powell, D. G., Radio Influence Voltage Characteristics of Transmission Line Assemblies, Using Semi-conducting Glazed Insulators, IEEE paper C73 416-5 presented at IEEE Summer Power Meeting, 1973, Vancouver, B.C.
48. Fukui, H., Naito, K., Irie, T., and Komoto, I., A Practical Study on Application of Semi-conducting Glaze Insulators to Transmission Line, IEEE paper T74 073-3, presented at IEEE Winter Power Meeting, New York, Jan. 1974.
49. Nigol, O., Reichman, J., and Rosenblatt, G., Development of New Semi-conductive Glaze Insulators, paper T73 420-7, *IEEE Trans. Power Appar. Syst.*, March/April, 1974, vol. PAS-93, pp. 614–622.
50. Bundled Circuit Design for 115-138 kV Compact Transmission Lines, EPRI Report EL 1314 (2 vols.), Feb. 1980.
51. Phase to Phase Switching Surge Design, EPRI Report EL 1550, Sept. 1980.
52. Souchereau et al., Validation of a Chainette Tower for a 735 kV Line, *CIGRE Paper* 22-04, 1978.
53. IEEE Committee Report: Limitations on Stringing and Sagging Conductors, *IEEE Trans. Power Appar. Syst.*, Dec. 1964, vol. 83, no. 12, pp. 1230–1235.
54. Converti, V., Hyland, E. J., and Tickle, D. E., Optimized Transmission Tower Spotting on Digital Computer, *AIEE CP60-1201*, Oct. 1960.
55. Peyrot, A. H., Peyrot, E. M., and Carton, T., Interaction and Integration in Power Line Design, *IEEE Computer Appl. Power*, 1992, vol. 5, no. 4.
56. Carton, T., and Peyrot, A. H., Computer-Aided Structural and Geometric Design of Power Lines, *IEEE Trans. Power Syst.*, Feb. 1992, vol. 7, no. 1, pp. 438–443.
57. ASCE Manual 74: *Guidelines for Electrical Transmission Line Structural Loading*, American Society of Civil Engineers, New York, 1991.
58. ASCE Standard 7-88: *Minimum Design Loads for Buildings and Other Structures* (revision of ANSI A58.1-1982), American Society of Civil Engineers, New York, 1990.
59. Ehrenburg, D. O., Transmission Line Catenary Calculations, AIEE paper, Committee on Power Transmission & Distribution, July 1935.
60. Winkelman, P. F., Sag-Tension Computations and Field Measurements of Bonneville Power Administration, AIEE paper 59-900, June 1959.

61. Fink, D. G., and Beaty, H. Wayne, *Standard Handbook for Electrical Engineers*, 13th ed., McGraw Hill, New York, 1993.

62. Aluminum Company of America, *Graphic Method for Sag Tension Calculations for ACSR and Other Conductors*, 1997.

63. Batterman, R. M., *ALCOA's Computer Program for Cable Sag and Tension Calculations*, Aluminum Company of America, Pittsburgh, Pa., 1967.

64. *Aluminum Association, Stress-Strain-Creep Curves for Aluminum Overhead Electrical Conductors*, publication SSCC-723613, The Aluminum Association., Inc., Waldorf, Md., 1997.

65. Greisser, V. H., Effects of Ice Loading on Transmission Lines, *Trans. AIEE*, 1913, vol. 32, p. 1829.

66. Healy, E. S., and Wright, J. A., Unbalanced Conductor Tensions, *Trans. AIEE*, 1926, p. 1064.

67. Poffenberger, J. C., and Swart, R. L., Differential Displacement and Dynamic Conductor Strain, *IEEE Trans.*, vol. PAS-84, 1965, pp. 281–289.

68. *Transmission Line Reference Book—Wind Induced Conductor Motion*, Electric Power Research Institute, Palo Alto, Calif., 1979, chap. 3.4.

69. Ibid., chap. 3.5.

70. Nigol, O., and Havard, D. G., Control of Torsionally-Induced Conductor Galloping with Detuning Pendulums, IEEE paper A78 125-7, Jan. 1978.

71. Richardson, A. S., Design and Performance of an Aerodynamic Anti-Galloping Device, IEEE Conf. paper C68 670-PWR, June 1968.

72. Douglass, D. A., and Roche, J. B., Anti-Galloping Potential of a New Twisted Conductor Design, *Proc. Canadian Electrical Association Int. Symp. Overhead Conductor Dynamics*, June 1981, pp. 83–98.

73. *Transmission Line Reference Book—Wind Induced Conductor Motion*, chap. 5.

74. Sellers, A. H., and Williams, J. E., All-Aluminum Transmission Tower Line, *Trans. AIEE*, June 1961, p. 169.

75. Thomas, Percy H., Formula for Minimum Horizontal Spacing, *Trans. AIEE*, 1928, vol. 47, p. 1323.

76. Farr, F. W., Ferguson, C. M., McMurtrie, N. J., Steiner, J. R., White, H. B., and Zobel, E. S., A Guide to Transmission Structure Design Loadings, *Trans. IEEE Power Group*, Nov. 1964, p. 1073.

77. Farmer, G. E., The Use of Insulated Ground Wires on a Transmission Line for Communication Purposes, *IEEE Trans. Power Appar. Syst.*, Dec. 1963, p. 884.

78. Ramthun, M. K., Pitzel, B. H., and Campbell, D. W., Stream-Lined 230-kV Transmission Passes Overhead in City's Streets, *Electr. World*, June 29, 1964, p. 94.

79. Stumpf, M. W., and Mouton, R. A., 12 Sided Single Poles Carry 760 MVA Capacity Line (New Orleans, La.), *Electr. World*, Nov. 16, 1964, p. 94.

80. ASCE Manual 52, *Guide for the Design of Steel Transmission Towers*, 2d ed., American Society of Civil Engineers, New York, 1988.

81. Richardson, W. B., New Techniques Speed Construction of 500 kV Lines, *Electr. World*, Jan. 15, 1965, p. 27.

82. Newmark, N. M., and Rosenblueth, E. H., *Fundamentals of Earthquake Engineering*, Englewood Cliffs, N.J., Prentice-Hall, 1971.

83. Klopfenstein, A., McDonald, J. F., Pecknold, D. A. W., and Walker, W. H., Seismic Test and Analysis of Capacitor Banks, *IEEE Trans. paper* T74 406-5, Jan./Feb. 1975, vol. PAS-94, no. 1, p. 81.

84. Long, L. W., Analysis of Seismic Effects on Transmission Structures, *IEEE Trans. paper* T-73-326-6, Jan./Feb. 1974, vol. PAS-93, no. 1, pp. 248–254.

85. Standard Method for Analysis of Creosote and Oil-type Preservatives, AWPA A1-98, Granbury, Tex., American Wood-Preservers Association., (http://www.awpa.com/)

86. Weber, L. C., Glass, E. C., and Alexander, G. W., Application of Statistical Methods in the Design and Uprating of Wood-pole Transmission Lines, *Trans. IEEE Power Group*, *Aug.* 1965, p. 725.

87. I*nstallation of Overhead Line Conductors*, IEEE 524, 1980.

88. Fowle, Frank F., The Transposition of Electrical Conductors, *Trans. AIEE,* 1904, vol. 23, p. 659.

89. Von Voigtlander, F., Transposition Practices, *Elec. Eng.,* Jan. 1943.

90. Clem, J. E., Currents Required to Remove Conductor "Sleet," *Electr. World*, Dec. 6, 1930, p. 1053, and Jan. 31, 1931, p. 245.

91. Live Line Maintenance Methods, *IEEE Trans. paper* T 73-157-5, *IEEE Trans. Power Appar. Syst.*, Sept./Oct., 1973, vol. PAS 92, no. 5, pp. 1642–1648.

92. Bibliography of Literature for Live-line Maintenance and Related Topics, *IEEE Trans. on Power Delivery*, July 1992, vol. 7, no. 3, pp. 1552–1562.
93. Doble, F. C., Progress in Field Testing of Insulators, *Electr. World*, 1923, vol. 81, p. 1397.
94. White, H. B., Design of Chute des Passes 345-kV Transmission Line, AIEE paper 60-72 WGM, 1960.
95. White, H. B., Cross Suspension System, Kemano Kitimat Transmission Line, AIEE paper CP 58-432 WGM, 1958.
96. Ritky, F., and White, H. B., Unique Suspension System Conquers Rugged Terrain, T & D World, Aug. 1997, vol. 49, no. 8.
97. White, H. B., Structural System for the James Bay Transmission Lines: Hydro Quebec Symposium on EHV and UHV Alternating Current, IREQ, 1973.
98. Lecomte, D., and Meyere, P., Evolution of the Design for the 735-kV Transmission Lines of Hydro-Quebec, CIGRE paper 22-08, 1980.
99. Cornell University, *Transmission Line Structure Foundations for Uplift/Compression Loadings*, Electric Power Research Institute Report EL-2870, Palo Alto, Calif., Feb. 1983.
100. Cornell University, *Critical Evaluation of Design Methods for Foundations under Axial Uplift and Compression Loads*, Electric Power Research Institute, Report EL-3771, Palo Alto, Calif., Nov. 1984.
101. GAI Consultants, Inc., *Laterally Loaded Drilled Pier Research*, vols. 1 and 2, Electric Power Research Institute, Report EL-2197, Palo Alto, Calif, Jan. 1982.
102. GAI Consultants, Inc., *Direct Embedment Foundation Research*, Electric Power Research Institute, Report EL-6309, Palo Alto, Calif., April 1989.
103. GAI Consultants, Inc., *Direct Embedment Foundation Research, Load Test Summaries*, Electric Power Research Institute, Report EL-6849, Palo Alto, Calif., June 1990.
104. Goldberg, D. T., Jaworski, W. E., and Gordon, M. D., *Lateral Support System and Underpinning*, vol. I, *Design and Construction*, prepared for Federal Highway Administration, U.S. Department of Commerce Publication PB-257 210, April 1, 1976.
105. Post-Tensioning Institute., *Post-Tensioning Manual*, 5th ed., Phoenix, Ariz., 1990.
106. Rural Utilities Services Bulletin 1724E-203, Guide for Undergrounding RUS Transmission Lines, 1994.
107. Southwire Company, Overhead Conductor Manual, 1994.
108. *Transmission Line Reference Book—115/138 kV Compact Line Design*, Electric Power Research Institute, 1978.
109. Barthold, L. O., and Barnes, H. C., High Phase Order Power Transmission; *Electra* 1973, no. 24, pp. 139–153.
110. Venkata, S. S., Guyker, W. C., et al., 138 kV Six Phase Transmission System: Fault Analysis, IEEE paper 81SM485-2, Summer Power Meeting, July 1981.
111. Stewart, J. R., Oppel, L. J., Thomann, G. C., Dorazio, T. F., and Brown, M. T., Insulation Coordination, Environmental and System Analysis of Existing Double Circuit Line Reconfigured to Six Phase Operation, *IEEE Trans. Power Delivery*, July 1992, vol. 7, no. 3, pp. 1628–1633.
112. Stewart, J. R., Oppel, L. J., Thomann, G. C., Dorazio, T. F., and Rebbapragada, R. V., Transformer Winding Selection Associated with Reconfiguration of Existing Double Circuit Line to Six Phase Operation, *IEEE Trans. Power Delivery*, April 1992, vol. 7, no. 2, pp. 979–985.
113. *Electr. World*, July 15, 1976 p. 63.
114. Apostolov, A. P., and Raffensperger, R. G., Relay Protection Operation for Faults on NYSEG's Six-Phase Transmission Line, *IEEE Trans. Power Delivery*, Jan. 1996, vol. 11, no. 1, p. 191.
115. Stewart, J. R., Oppel, L. J., and Richeda, R. J., Corona and Field Effects Experience on an Operating Utility Six-Phase Transmission Line, *IEEE Trans. Power Delivery*, Oct. 1998, vol. 13, no. 4, pp. 1363–1369.
116. Landers, T. L., Richeda, R. J., Krizauskas, E., Stewart, J. R., and Brown, R. A., High Phase Order Economics: Constructing a New Transmission Line, IEEE paper PE-405-PWRD-0-12-1997 presented at the IEEE Power Engineering Society, 1998 Winter Meeting, Tampa, Fla.
117. Oppel, L. J., Krizauskas, E., and Austenfeld, R. H., Evaluation of the Performance of Line Protection Schemes on the NYSEG Six Phase Transmission System, *IEEE Trans. Power Delivery*, Oct. 1998, vol. 13, no. 4, pp. 1521–1526.
118. Brown, R., Landers, T., Stewart, J., and Oppel, L., Six-Phase Successfully Applied to Utility Transmission System, *CIGRE paper 22/33/36-01*, Session 1998, Paris.

14.2 UNDERGROUND POWER TRANSMISSION

By E. C. BASCOM, III and J. A. WILLIAMS

14.2.1 Cable Applications

Traditionally, underground cable systems have been installed in major urban areas where overhead lines are not practical—locations such as airport approaches because of safety issues or water crossings. Cables are generally much more costly than overhead lines from the standpoint of material and installation costs, although recent trends regarding rights-of-way and permitting costs often make the underground alternative more competitive. Increasingly, cables are being selected because existing rights-of-way are too congested or unavailable, or local municipalities will not tolerate a new overhead transmission line. Transmission cables may also be applied for substation getaways, crossings under major overhead line corridors, directionally drilled installations and other installations where technical considerations favor underground cables. Utilities sometimes find it prudent to place short dips in critical areas of overhead circuits, to allow installing overhead lines, which would not otherwise be permitted.[1]

Extruded dielectric (XD)—principally cross-linked polyethylene (XLPE)—cables have become the U.S. standard for voltages up to 230 kV, with short installations up to 345 kV. Over the last 7–8 years, XLPE cables have been used worldwide at 400 and 500 kV. High-pressure fluid-filled pipe-type cables are less commonly installed because of concerns about the dielectric liquids and higher maintenance, but there are a few thousand circuit miles that continue to operate, and several new circuits are installed annually. New self-contained liquid-filled (SCLF) cable installations are being displaced by the availability of XLPE cables at higher voltages. Mass-impregnated nondraining (MIND) paper cables are used for high voltage direct current (HVDC) circuits, although one manufacturer has started supplying polymeric-insulated cable for dc applications. Compressed-gas-insulated transmission systems are uncommon for new installations and used only for special applications—typically short lengths and high-power transfers within substations. In the United States, even though most new installations are extruded dielectric, pipe-type cables account for about 70% of the approximately 6000 circuit miles (10,000 km) in service; XD accounts for 25%, and self-contained liquid-filled accounts for the remainder—with a very small amount of compressed-gas-insulated cables.

14.2.2 Cable System Considerations and Types

Cable Integration into Utility System. Planning and operating considerations for underground cables are different from those for overhead lines.[2] Special attention should be paid to properly represent the cables for utility system analyses, capacitance effects including charging currents, reactor application, system restoration, inductance effects including load sharing, surge impedance loading, insulation coordination, system insulation requirements, and losses. As cables are used at higher voltages—230 kV and above—charging current and reactive compensation should be carefully evaluated. In addition to the effect of cable capacitance on system operation, the current required to charge the capacitor reduces the allowable cable rating. Charging current can be determined from Eq. 14-89, which shows that the charging current increases proportionally with voltage.

$$I_{Charging} = 2\pi f C E_0 = \frac{2\pi f \varepsilon E_0}{18 \ln \frac{D_{INSULATION}}{D_{CONDUCTOR}}} \times 10^{-9} \tag{14-89}$$

In Eq. 14-89, f is the power frequency, C is the capacitance, ε is the insulation dielectric constant (specific inductive capacitance), and E_0 is the line-to-ground voltage in volts.

Table 14.24 lists several characteristics of overhead lines and underground cables.[3] Cables tend to "hog" load compared to overhead lines because of the lower surge impedance as compared to

TABLE 14-24 Typical Electrical Characteristics, 230-kV Overhead Line and Underground Cables

Parameter	Overhead Line	Underground XLPE	Underground HPFF (PPP)
Shunt capacitance, μf/F/mi	0.015	0.30	0.61
Series inductance, mH/mi	2.0	0.95	0.59
Series reactance, ohm/mi	0.77	0.36	0.22
Charging current, A/mi	1.4	15.2	30.3
Dielectric loss, kW/mi	0+	0.2	2.9
Reactive charging power, MVA/mi	0.3	6.1	12.1
Capacitive energy kJ/mi	0.26	2.3	7.6
Surge impedance, ohms	375	26.8	14.6
Surge impedance loading limit, MW	141	1975	3623

overhead lines, although they generally have a lower thermal rating than comparable overhead lines. Since cable ratings are usually lower, potential load flow problems should be investigated when integrating long cable lines in parallel with overhead lines.

14.2.3 Extruded-Dielectric Systems

Extruded-dielectric systems sometimes called solid-dielectric cables provide a simpler, often lower-cost alternative to the paper-insulated cables that have historically been used for transmission cables. Extruded-dielectric cable types have included linear low-density polyethylene (LLDPE), cross-linked polyethylene (XLPE), and ethylene-propylene-rubber (EPR), although XLPE is now the most common insulation type used for transmission cables particularly since the challenges of manufacturing EHV cables with XLPE insulation have been addressed. The absence of dielectric fluid greatly simplifies the ancillary equipment and accessory complexity, and removes concerns about fluid leaks. The XD cables also have lower capacitance than paper-insulated cables, simplifying their integration into the utility system. Although the first XD cables were installed in the United States in the 1960s, they did not find extensive use until the mid 1980s.

Extruded-dielectric cables consist of a copper or aluminum conductor, a conductor semiconducting shield to ensure a smooth electrical profile in the insulation, an extruded polymeric insulation, an extruded insulation shield, outer metallic shielding, and typically a moisture barrier such as a lead sheath or metallic foil laminate, covered with a polymeric (usually polyethylene) jacket (Fig 14-73).

FIGURE 14-73 345-kV extruded-dielectric cables. (Courtesy of LS Cable.)

The "true triple extrusion" method is the state-of-the art method for applying the semiconducting shields and insulation in one pass through the extrusion line to provide sound, void-free interfaces, especially with transmission cables. Extruded-dielectric cable insulation is extremely sensitive to the presence of partial discharge; consequently, this cable must operate under ionization-free conditions.

Exacting manufacturing standards supported by an effective quality assurance system and extensive factory testing are all necessary to provide a high-quality product suitable for application at the highest transmission voltages.

Contrary to initial industry expectations, XD insulation has proven to be sensitive to moisture so transmission cables use moisture barriers to prevent ingress into the insulation. Water blocking is sometimes used on stranded conductors, and water-swellable textile tapes or powder are applied under the metallic sheath. Early cables were steam cured, but most high-voltage extruded-dielectric cables manufactured since the mid-1990s are dry-cured and many are dry cooled (CDCC, continuous dry curing and cooling). Moisture-barrier conductors and swellable tapes under the sheath help maintain the cable free from moisture, particularly when the cable sustains mechanical damage such as from a dig-in. Great improvements in the cables and their accessories in the last decade have led the U.S. utilities to conclude that extruded-dielectric cable reliability is now comparable to paper cables.

Sheath bonding is an important design consideration and requires more careful consideration than at distribution voltage levels. If both ends of the sheath are grounded as is common in distribution, conductor current induces currents in the sheaths, causing heat losses that derate the cable by 10–35%. If the sheath is grounded at one point, the heat losses from circulating currents are eliminated but sheath voltages are induced, equal to approximately 150 V/1000 A/km. Cross bonding, electrically transposing the sheaths, can eliminate sheath circulating currents (although eddy currents still remain) while maintaining acceptable induced voltages.[4]

Extruded-dielectric cables are typically installed individually, either directly buried or in duct. In some cases, the three individual phases are installed at one time in a plastic or steel pipe as part of a “trenchless” installation, in a manner resembling a pipe-type cable installation. Splice spacing ranges from 300 to 1000 m, depending upon the installation mode, reel-shipping lengths, and allowable sheath voltages.

Taped, taped- and field-molded, prefabricated, and premolded splices (Fig. 14-74) have been used; although, premolded splices (joints) are used almost exclusively today.

Terminations have either porcelain or polymer housings, and most designs have a small amount of dielectric liquid (e.g., silicone oil) or gas (SF_6) in the termination.

FIGURE 14-74 “Click-fit” premolded joint. (Courtesy of Prysmian.)

14.2.4 High-Pressure Fluid-Filled (HPFF) Systems

Pipe-type cable had been the U.S. standard because of its ruggedness, and the relative ease of installing pipe in city streets. The first pipe-type cable was a 66-kV circuit installed in Philadelphia in 1932. Many hundreds of miles of 230- and 345-kV pipe-type cables were installed in the 1960s and 1970s. Its use was declined in the late 1980s, as more extruded-dielectric cables were installed and as heavy voltage cable usage generally diminished.

A welded, coated, cathodically protected steel pipe, typically 8.625 or 10.75 in optical density, is pressure and vacuum tested, and the three mass impregnated cables are pulled together into the pipe (Fig 14-75). The cables consist of a copper (occasionally aluminum) conductor, conductor shield, taped insulation, insulation shield, outer shielding, and skid wire to prevent cable damage as they are pulled into the pipe. High-quality kraft paper is typically used at 138 kV and sometimes 230 kV. Introduction of a laminated paper—polypropylene—sometimes referred to as PPP for paper-polypropylene-paper—insulation in the 1980s permitted a lower insulation thickness (0.600 in at 345 kV versus 0.920 in for kraft paper), with lower electrical losses, so this type of insulation is now selected for most new 230 kV and all 345 kV circuits.

FIGURE 14-75 Pipe-type cable.

Splices are typically spaced at 2000–3500 ft intervals. Terminations with porcelain housings make the transition to air insulation and atmospheric pressure.

For a high-pressure liquid-filled (HPLF) system, which has been proven experimentally up to 765 kV though never commercially applied above 345 kV in the United States, the pipe is pressurized to approximately 200 psig (1.4 MPa) with dielectric liquid to suppress ionization. Expansion and contraction of the liquid requires large reservoir tanks and sophisticated pressurizing systems. This provides the possibility of removing the fluid periodically along the length of the line, cooling it (called “forced cooling”), and returning it to the pipe to obtain a 20–50% ampacity increase. Pipe-type cables offer several opportunities for uprating because of the presence of the dielectric liquid.

Nitrogen gas at 200 psig (1.4 MPa) may be used to pressurize the cable at voltages to 138 kV. This high-pressure gas-filled (HPGF) system requires a slightly greater cable insulation thickness because of the poorer insulating qualities of the gas. However, the system greatly reduces environmental concerns because there is limited free dielectric liquid if there is a leak in the pipe, and no external pressurizing equipment is required other than a regulator, nitrogen cylinder, and alarms.

14.2.5 Self-Contained Liquid-Filled (SCLF) Systems

This system was developed in the 1920s, and was a worldwide standard outside the United States through the 1980s. There are many miles in operation at 525 kV, for both land and submarine cables. The cable conductors have a hollow core that provides a longitudinal feeding path for the dielectric liquid. A low-pressure design requires less than 10 psig (70 kPa) pressure to suppress ionization, while a medium-pressure design typically operates at 50–75 psig (350-515 kPa). The 525-kV cables at Grand Coulee dam have a reinforced sheath; the lower end of these cables operates at 400 psig (2.8MPa).[5]

The cables are maintained gas free at a positive pressure from manufacturing through energization and operation because the lower operating pressure (than pipe type) would allow gas ionization, if present. The liquid volume is much lower than that for an HPFF cable. Fluid expansion and contraction is accommodated by bellows-type expansion tanks distributed along the route at spacings

FIGURE 14-76 Self-contained cable. (Courtesy of Prysmian.)

that depend upon sheath construction, core diameter, type of liquid, temperatures, circuit topography, and power transfer. Special fluid stop joints are installed to divide the cable route into a number of hydraulic sections as required because of static pressures (due to route profile) or dynamic pressures (due to fluid volume changes during heating and cooling transients). For very long circuits such as submarine cables, an active pressurizing plant is employed, resembling the plant used for fluid-filled pipe-type cables.

Figure 14-76 shows a self-contained cable. The conductor is most commonly copper, although aluminum has been used. Conductor shields, insulation, and insulation shields are similar to those for pipe-type cables. A lead, reinforced lead, or aluminum sheath maintains pressure, excludes moisture, and provides a fault current path, although submarine cable circuits often include additional copper to support return currents. Aluminum sheaths are usually corrugated to improve flexibility. A polymeric jacket isolates the sheath from ground and provides mechanical and corrosion protection.

Self-contained cables are especially suited for long-distance water crossings because they can be produced in long splice-free lengths, which minimize—or in some cases, eliminate—the number of field splices. Wire armor is provided to aid in cable installation and retrieval, and give a small measure of mechanical protection.

14.2.6 Direct Current Cables

DC cables are used for long water crossings and their shore ends. Charging current and dielectric losses make ac cables unsuitable for lines more than 30–50 mi long (50–80 km), depending upon voltage level and insulation type. Traditional HVDC cables used a mass-impregnated solid-type cable impregnated with a very high-viscosity fluid that had no active fluid pressurizing system (Fig. 14-77). This cable is used for very long (>25 mi, >40km) dc submarine applications. Self-contained liquid-filled or pressurized gas-filled cables have been used for the shorter lengths. Table 14-25 provides a listing of major dc installations. Polymeric insulation materials have recently been developed that are suitable for dc applications, usually combined with IGBT (voltage source) valve converters instead of thyristor valves; these types of HVDC cables can feed smaller loads not connected to large ac systems as was traditionally required.

The insulation for dc cables is especially critical since the insulation is resistively graded instead of capacitively graded as with ac systems. The electrical stress distribution in a dc cable is a function of insulation temperature—electrical stresses can be higher at the outer shielding, whereas for ac cables, the stress is always highest at the conductor shielding. Mass-impregnated cables are generally limited to 50°C-conductor temperatures to avoid temperature distributions that would give too high a stress at the outer shielding.

14.2.7 Gas-Insulated Transmission Lines (GITL)

These "cable" types consist of sulfur hexafluoride (SF_6) gas or a mixtures of SF_6 and nitrogen at pressures from 30–45 psig (200–315kPa) serving as an insulation between a coaxially spaced tubular aluminum conductor and tubular aluminum shield. Insulating spacers maintain separations as shown in Fig. 14-78.

The system is supplied in rigid 40-ft lengths. Large diameters are required because the gas is a poorer insulator than paper-insulated or extruded-dielectric cables—the first 345-kV system had an

18-in enclosure diameter for each of the individual phases. The large diameter enhances heat transfer, conductor sizes are large, and the SF_6 insulation does not have the thermal limitations that other cable types have. Ampacities are therefore very high, usually able to match overhead line capacities with one conductor per phase so these cables often are employed to make high-capacity bus connections within substations. The system is especially suited for short distances, very high currents, and extra-high voltages—the first system was a 600-ft, 2000 A, 345-kV installation.[6]

FIGURE 14-77 HVDC mass-impregnated cable. (Courtesy of Prysmian.)

14.2.8 Superconducting Cables

Superconducting wire has existed for several decades, taking advantage of liquid helium (4 K) temperatures (“low temperature” superconductors) to cool various materials to the point that they show no appreciable electrical resistance. Within the last 5 years, “high temperature” superconductors (HTS) using liquid nitrogen temperatures (80 K) were developed along with methods to make the wire in length suitably long to apply to cables. Two basic types of HTS cable exist; one is the “cold-dielectric” type (Fig. 14-79) where the insulation is at cryogenic temperatures, while the other is a warm dielectric where only the conductor is operating at cold temperatures.

High-temerature superconductors-cable technology offers advantages since the HTS-cable system operates independent of the thermal environment in which it is installed while allowing significant power transfer—using high current densities—with single circuits potentially matching the capacity of overhead lines. HTS cables also offer the possibility of transferring bulk power at what have traditionally been medium voltages by taking advantage of the high current density capability rather than stepping up the voltage as is typically done with conventional “transmission” circuits. One possible application is to retrofit an HTS cable into existing cable pipes or conduits to provide for a significant power transfer increase without the high associated costs of construction and obtaining rights-of-way.

While there has been aggressive research to apply HTS cables at transmission voltages, they have seen only limited use in commercial power systems. Long-term reliability of the cryogenic cooling plants has not been proven, and accessories—joints and terminations—require further development. Cost is also a significant factor in making these systems commercially viable. Materials and installation costs for HTS cable systems are close to 20 times that of a conventional power cable system. Even at this price point, there are certain applications—usually where construction costs, permitting or obtaining rights-of-way are impractical—that HTS cables could see some use. The technology is expected to improve with new generations of superconducting wire becoming available and improved reliability of accessories and cooling plants.

14.2.9 Cable Capacity Ratings: Ampacity

The ampacity (current rating) for underground cables is limited by the maximum conductor temperature that can occur in the insulation with rated current flow through the conductor. The insulation material that is in contact with the conductor shield limits the maximum conductor temperature.

TABLE 14-25 DC Cable Installations

Name of Link	Date	Voltage (kV)	Power (MW)	Length (km)	Cable Type
Gotland 1	1954	100	20	100	MI
Cross Channel 1	1961	100	80	2 × 52	MI
SACOI	1965	200	100	2 × 118	MI
Cook Straight 1	1965	250	300	64	PIGF
Konti-Skan 1	1965	285	300	64	MI
Vancouver 1	1969	300	156	3 × 27	MI
Mallorca/Menorca	1972	200	100	3 × 44	SCFF
Skaggerak 1,2	1976	263	250	2 × 125	MI
Vancouver 2	1976	300	185	2 × 35	MI
Hokkaido/Honshu	1980	250	150	2 × 42	SCFF
Gotland 2,3	1983	150	160	2 × 100	MI
Cross Channel 2	1986	270	250	8 × 50	MI
Konti-Skan 2,3	1988	285	300	2 × 64	MI
Fenno-Skan	1989	400	500	200	MI
Cook Strait 2	1991	350	500	3 × 40	MI
Skagerrak 3	1993	350	500	125	MI
Cheju (Korea)	1993	180	150	2 × 96	MI
Baltic	1994	450	600	250	MI
Morocco-Spain	1995	300	600	26	SCFF
Kontek (Germany)	1995	400	600	170	MI
Gotland	1999	+/–80	65	2 × 70	Polymer
Kii Channel	1999	500		4 × 50	SCFF
Sweden-Poland	2000	450	600	235	MI
Moyle (NIE)	2000	250	500	2 × 55	MI
Italy-Greece	2001	400	500	200	MI
DirectLink	2001	+/–80	180	6 × 59	Polymer
Norway-Germany	2001	500	600	578	MI
MurrayLink (Australia)	2002	+/–150	200	2 × 180	Polymer
Cross Sound	2002	+/–150	330	2 × 40	Polymer

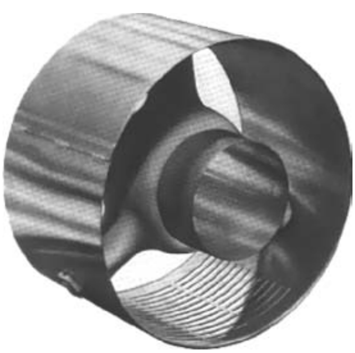


FIGURE 14-78 GITL spacer. (Courtesy of CGIT Westboro.)

The conductor temperature is determined by the heat generated within the cable—electrical losses (I^2R) and dielectric losses (ac cables only)—which passes radially to ambient earth through various thermal resistance layers in the cable, duct (when present), trench, and surrounding soil. The temperature of the conductor can be determined from the following thermal equivalent circuit (this circuit is shown for XD or super conducting final focus [SCFF] cables).

Ampacity can be calculated for XD, SCFF, and HPFF cables by solving the equivalent thermal circuit for conductor temperature based on the type of cable being considered. Figure 14-80 shows an equivalent thermal circuit for an extruded transmission cable installed in conduits. The earth thermal resistance depicted in Fig. 14-80 includes the concrete envelope or other special backfill, if present. Equation 14-90 for the XD cable modeled in Fig. 14-80 is:

$$T_{Conductor} = T_{Ambient} + W_{Dielectric} \times \left(\frac{\overline{R_{Insulation}}}{2} + \overline{R_{Jacket}} + \overline{R_{Jacket-to-Duct}} + \overline{R_{Duct}} + \overline{R_{Earth}}\right) + I^2R_{Conductor} \times \left(\overline{R_{Insulation}} + Qs(\overline{R_{Jacket}} + \overline{R_{Jacket-to-Duct}} + \overline{R_{Duct}} + \overline{R_{Earth}})\right) \tag{14-90}$$

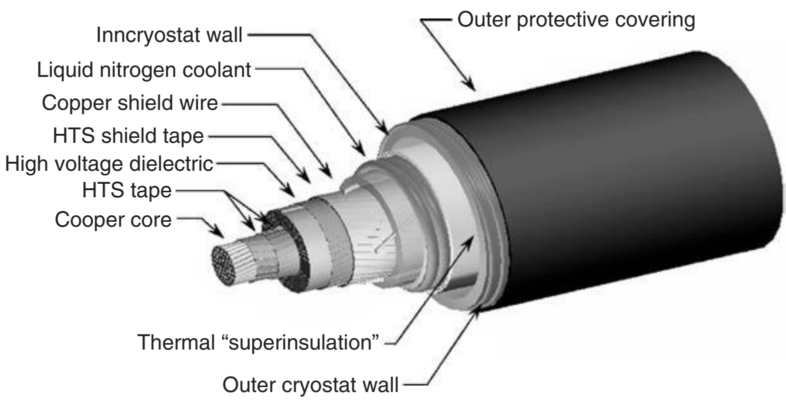


FIGURE 14-79 Cold dielectric HTS superconducting cable. (Courtesy of American Superconductor.)

where the various variables are noted as in Fig. 14-80 with *Qs* being the ratio of losses from the conductor and sheath to the losses only in the conductor.

The earth thermal resistance portion of the equivalent thermal circuit is defined by Eq. 14-91:

$$\overline{R_{earth}} = \frac{\overline{\rho_{backfill}}}{2\pi}\cdot n\cdot\left(ln\left(\frac{D_x}{D_{earth}}\right) + LF\cdot ln\left(\frac{2\cdot L + \sqrt{4\cdot L^2 - D_{earth}^2}}{D_x}\right)\right)$$
$$+\frac{\overline{\rho_{backfill}}}{2\pi} n\cdot LF\cdot ln(F) + \frac{\overline{\rho_{native}} - \overline{\rho_{backfill}}}{2\pi}\cdot n\cdot N\cdot LF\cdot G_b \qquad (14\text{-}91)$$

where $\overline{\rho_{backfill}}$ is the thermal resistivity of the backfill in °C-m/W, *Dx* is the diameter beyond which the average (rather than peak) daily losses are experienced, D_{earth} is the outer diameter of the cable, pipe or conduit, *n* is the number of cables within that earth diameter, *L* is the burial depth to the center of the backfill, *LF* is the daily loss factor (ratio of peak losses to average losses), *F* is a mutual heating effect factor for *N*-1 cables or pipes with *N* being the number of cables/pipes/conduits installed, $\overline{\rho_{native}}$ is the native soil thermal resistivity in °C-m/W and *Gb* is a geometric factor for the backfill envelope.

For ampacity calculations, the dielectric heating is assumed to be constant since the voltage level is generally held to be constant. Also, we can calculate the temperature rise from ac heating as

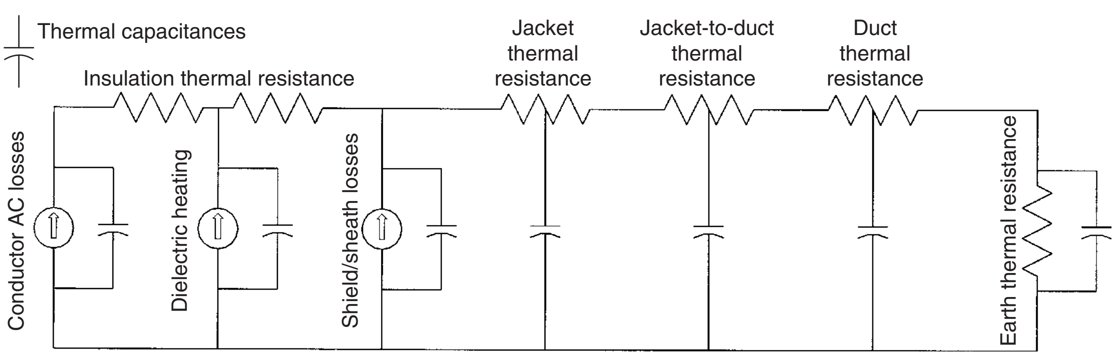


FIGURE 14-80 Equivalent thermal circuit for XD cable installed in conduit.

$I^2 \times (\Sigma R_{AC} \times \overline{R_{Thermal}})$. Then the conductor temperature can be found as $T_{conductor} = T_{Ambient} + \Delta T_{Dielectric} + I^2 R_{AC} \times \overline{R_{Thermal}}$ from which the current for a given conductor temperature can be found:

$$I = \sqrt{\frac{T_{Conductor} - \Delta T_{Dielectric} - T_{Ambient}}{\Sigma R_{AC} \times \overline{R_{Thermal}}}} \tag{14-92}$$

A detailed procedure for calculating ampacity is beyond the scope of this reference. The reader is referred to Refs. 7 and 8 for additional information. The main elements that make up the thermal circuit include the following:

Conductor resistance, is determined based on the cross-sectional area of the conductor, the conductor material (copper or aluminum), the maximum operating temperature, the ac skin, and proximity effects. The maximum operating temperature is defined by the insulation material. Industry-accepted limits for the maximum conductor temperature are 85°C for paper or laminated paper-polypropylene insulations, 75°C for high-density polyethylene insulation, and 90°C for cross-linked polyethylene or ethylene-propylene-rubber insulations. Following the Association of Edison Illuminating Companies (AEIC) practices, these maximum temperatures are reduced by 10°C when the earth thermal parameters are not well defined.

Insulation thermal resistance is expressed in K-m/W (1 K temperature rise occurs when 1 W of heat flows through a thermal resistance of 1 K-m/W). The thermal resistance of the insulation or any cylindrical layer of cable material (i.e., jacket, pipe coating, and conduit) with an inner diameter D_i, outer diameter D_o, and material thermal resistivity of $\overline{\rho}$ can be found as follows:

$$\overline{R_{insulation}} = \frac{\overline{\rho}}{2\pi} ln\left(\frac{D_o}{D_i}\right) \tag{14-93}$$

Typical values for various cable materials are shown in Table 14-26.

Dielectric loss is the heat generated in ac cable insulation as a function of diameters, voltage, insulation temperature (indirectly, as a function of temperature-dependent dissipation factor), specific inductive capacitance, and dissipation factor. Reference 7 should be used to calculate dielectric losses.

Shield/sheath losses are conductor current-dependent losses that result from induced circulating and eddy currents in the cable shield and sheath. Their value depends strongly on the sheath bonding method as described in Sec. 14.2.2.

Earth thermal resistance depends on the cable burial depth, cable phase spacing, cable diameter, native soil thermal resistivity and special backfill thermal resistivity. For a single cable buried in the soil, the thermal resistance can be found from Eq. 14-94:

$$\overline{R_{earth}} = \frac{\overline{\rho_{earth}}}{2\pi} ln\left(\frac{2L + \sqrt{4L^2 - D_e^2}}{D_e}\right) \tag{14-94}$$

L is the burial depth to the center of the cable, and D_e is the diameter to the outside of the cable (for direct buried cables) or the outside diameter of the conduit (for cables in conduit). Transmission circuits generally have one cable for each of the three phases, and some installations

TABLE 14-26 Typical Values for Various Cable Materials

Material	Thermal Resistivity, $\overline{\rho}$ K-m/W
High-density polyethylene (HDPE)	3.5–4.0
Cross-linked polyethylene (XLPE)	3.5
Ethylene propylene rubber (EPR)	4.0–5.5
Oil/paper/laminated paper-polypropylene	5.0–6.0
Polyvinyl chloride (PVC)	4.0–7.0

TABLE 14-27 Typical Values of Soil Thermal Resistivities

Material	Thermal Resistivity, $\bar{\rho}$K-m/W
Fluidized Thermal Backfill (FTB)	0.40–0.75
Concrete	0.30–0.80
Stone screenings	0.40–1.00
Thermal sand	0.50–1.00
Uniform sand	0.70–2.00
Clay	1.00–2.50
Soil with high-organic content	2.00–4.00

include multiple circuits within the same trench. To account for mutual heating that can occur, an additional term must be included in the total earth thermal resistance, as follows:

$$\overline{R_{mutual}} = \frac{\overline{\rho_{earth}}}{2\pi} ln\left(\frac{d_{i1}'}{d_{i1}} \cdot \frac{d_{i2}'}{d_{i2}} \cdot \; \cdots \; \cdot \frac{d_{in}'}{d_{in}}\right) \tag{14-95}$$

where d_{in}' is the distance from cable i to the image of cable n, and d_{in} is the actual distance from cable i to n. For a single 3-phase circuit, $n = 2$. For a double 3-phase circuit (6 cables), $n = 5$. The hottest cable should be used as cable 1.

The soil thermal resistivity varies from 0.5 to 4.0 K°-m/W and is dependent on moisture content, material, mixture, size of particles, and degree of compaction. The earth thermal resistance accounts for up to 75% of the total thermal resistance from conductor to ambient. Although it is not possible to control the native soil thermal resistivity, it is common to backfill the cable trench with special low-resistivity materials such as thermal sand or Fluidized Thermal Backfill (FTB). Most extruded and self-contained liquid-filled systems in duct are encased in a high-strength concrete envelope for mechanical protection and to promote heat transfer. Reference 8 describes the way to model the thermal resistance of this envelope of material. Typical values of soil thermal resistivities are listed in Table 14-27.

Thermal capacitances shown in the thermal circuit are important for calculating the temperature response to load cycle changes or emergency ampacity capability. References [9] and [10] discuss the thermal capacitances of cables in detail. When load changes occur on an underground cable circuit, the temperature of the conductor does not respond instantaneously. A simplified load shape—rectangular load pattern—applied to a previously unloaded extruded-dielectric cable circuit is shown in Fig. 14-81. Note that the hourly fluctuations in the load pattern are in close correlation with the temperature response; these changes reflect the relatively short thermal time constant—a few

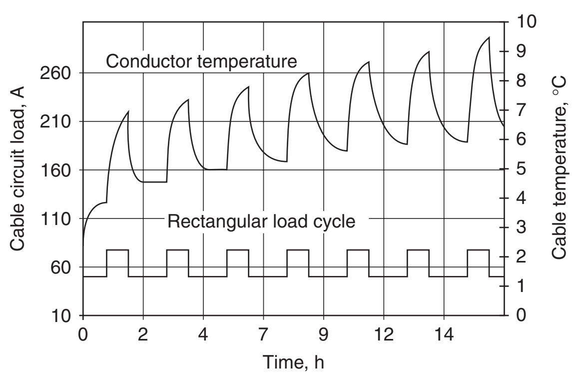


FIGURE 14-81 Temperature response to step load shape.

hours—for cable components. In contrast, the overall average temperature is rising as the duration of the applied load increases; this reflects the relatively long thermal time constant of the earth—approximately 50–150 h—for the mass of soil around the cables.

The impact of the earth thermal capacitance is routinely included in ampacity calculations by using a daily (24 h) "loss factor." The loss factor is essentially the load factor of the losses and is the ratio of the peak losses to the average losses over a day according to Eqs.14-96 and 14-97. Beyond a certain diameter out in the earth, the average effects of heat loss, rather than the peak losses, are experienced. So, the thermal resistances are adjusted accordingly to consider this effect. As the loss factor drops, the ampacity increases. The loss factor concept was initially developed in the Neher-McGrath paper in October, 1957 and is used by many utilities, particularly in North America. International practice (IEC-60287) has been used to calculate ampacities with 100% loss factor (flat load shape), which is generally very conservative since most cable circuits experience at least some load cycling over a 24-h period. Advanced numerical techniques have been applied on a limited basis to consider longer load cycles (several days to several weeks), but the ability to predict load over such long periods limits their applicability.

The mechanics of calculating ampacity are not particularly difficult, but there are many details to be considered. All of the calculations can be done using a hand calculator but are amenable to straightforward computer programming or a computer spread sheet. Ratings for many standard configurations are tabulated in Ref.[11].

High-Pressure Fluid-Filled Cables. Ampacity for HPFF cables may be calculated in a similar manner as for single-core cables. The main difference is that there is an additional element—the cable pipe—that generates heat in the equivalent thermal circuit. The pipe itself heats, and the magnetic pipe causes additional ac losses in the cables. A subtle difference for HPFF cables is that the heat generated by three cables—the three cable phases—must pass through the dielectric fluid in the cable pipe, the pipe coating, and the earth.

Compressed-Gas-Insulated Systems. Heat loss from GITL is by convection and radiation through the air since these systems are almost exclusively installed in air; usually open air but sometimes in enclosed troughs or tunnels. Consequentially, the rating procedure is significantly different from buried cable systems where heat loss is by conduction through the soil. The procedure for calculating ratings is generally an iterative process needed to evaluate the fourth-order relationships—guessing a temperature and rating and then comparing the calculated radiation and convection heat loss to the ohmic (I^2R) heat generated by the bus.

$$W_{radiation} = 5.69 \times 10^{-8}(T_2^4 - T_1^4)\pi D\left(\frac{1}{\frac{1}{e} + \frac{1}{\varepsilon} - 1}\right) \tag{14-96}$$

$$W_{convection} = 8.523\pi D\sqrt{\frac{(T_2 - T_1)^4}{\frac{T_2 - T_1}{2}}} \tag{14-97}$$

With T_2 and T_1 representing the temperature of the GITL surface and ambient air, respectively, D representing the outer diameter of the emitting surface, $e = [1\text{-}D/(6\pi d)]^2$, and ε representing the emissivity of the emitting surface. Details of the calculations can be found in Refs. 12 and 13. For the situation where the GITL is exposed to solar radiation, the methodology used to model solar radiation on overhead transmission lines can be applied.[7, 14, 15]

Transient and Emergency Ratings. Emergency ampacity considers the use of a higher operating temperature for a period combined with the heat storage capacity of the cable and earth environment. The temperature response of a single cable to a given heat output, W, can be found from Eq.14-98:

$$\theta_{Earth}(t) = W_{Cable}\frac{\overline{\rho_s}}{4\pi}\left[-Ei\left(-\frac{D_{earth}^2}{16\delta t}\right) + Ei\left(-\frac{L^2}{\delta t}\right)\right] \tag{14-98}$$

TABLE 14-28 Acceptable Operating Temperature Values

Material	Maximum Temperature (Normal), °C	Maximum Temperature (Emergency), °C
High-density polyethylene (HDPE)	75	90 (<1500 h over cable life)
Cross-linked polyethylene (XLPE)	90	105–130 (<72h)
Ethylene propylene rubber (EPR)	90	130 (<1500 h over cable life)
(Oil/paper/laminated paper Polypropylene	85	100 (<300 h)–105 (<100 h)

In compliance with AEIC specifications, these temperatures must be reduced by 10°C if the thermal environment is not defined for the entire circuit. A thermal route survey, ideally done prior to cable installation, eliminates this requirement. Dielectric fluid circulation in HPFF cables will also help mitigate the effects of localized conditions.

where δ is the earth thermal diffusivity in mm^2/h t is the time of the transient in seconds, L is the burial depth to the center of the cable, conduit, or pipe D_{earth} is the diameter of the cable/conduit/pipe-to-earth interface and ρ_s (needs "overline") is the thermal resistivity of the soil in K°-m/W and *Ei* is a notation to indicate an exponential integral. Further details for calculating emergency ratings may be found in Refs. 10 and 16.

Transient ratings are those in which the heat-storage capacity of the cable system permits short-duration overcurrents without exceeding allowable temperatures. The maximum allowable temperatures and permissible emergency rating durations are defined by cable specifications and agreement with manufacturers. Generally, acceptable operating temperature values are listed in Table 14-28.

A detailed procedure for calculating emergency ratings is available in Refs. 9 and 10.

Forced Cooling. Forced cooling is not very common and is generally considered only for HPFF cables where the dielectric liquid in the steel pipe can be circulated. There are some limited applications where water is circulated through parallel water pipes to provide forced cooling of extruded or self-contained cables. In HPFF cable systems, the dielectric liquid can be removed from the pipe, cooled by various types of heat exchangers, and reintroduced into the pipe at a remote location, generally using a noncabled liquid supply pipe (note that the heat capacity of nitrogen is small so HPGF lines are not forced cooled). The procedure can increase the capacity of the cable circuit by providing a lower-resistance path for the heat generated in the cable to escape to ambient. Figure 14-82 illustrates a typical hydraulic circuit for a forced-cooled cable.

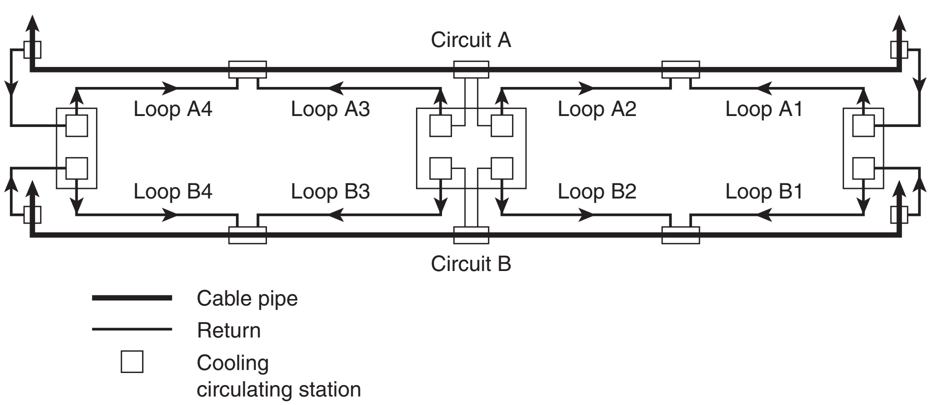


FIGURE 14-82 Hydraulic circuit, typical forced-cooled HPFF cable system.

The major additional components of a forced-cooled HPFF system include a return pipe for the fluid, heat exchangers, circulating pump (800–1600 L/min), lower-viscosity dielectric fluid, and modifications to the cable pipe to prevent cable damage from impinging oil. Reference 17 describes a typical 345 kV forced-cooled system, Ref. 18 describes a procedure for making thermal calculations, and Ref. 19 describes full thermal and hydraulic calculation procedures for pipe-type cables.

The added cost and complexity of forced-cooling equipment, along with potentially high-energy charges for operation of the equipment, have tended to limit forced-cooled installation to areas of very high-installation costs, or to retrofitting on existing HPFF cable lines.

Self-contained and extruded cables can be forced-cooled [18] using parallel water pipes (lateral cooling), conduit water circulation (integral cooling), heat pipes, or internal (core) cooling (Fig. 14-83). One North American utility uses parallel water-cooling pipes to increase the ampacity of a submarine cable as it comes on shore; the shore zone for submarine cables is often limiting, so the cooling system mitigates the ampacity limitation.

Internal cooling for self-contained liquid-filled cables has the advantage of removing dielectric fluid with high heat content from the core of the cable itself. This system requires special joints to periodically remove the line-potential liquid, pass it through heat exchangers, and return it to the cable core.

Lateral and integral cooling have a somewhat lower efficiency, but the simplicity of the cooling system has lead to several commercial installations. Extruded-dielectric cables can make use of either integral or lateral cooling. It is also possible to cool CGIT in the same manner, although there have been no such installations to date.

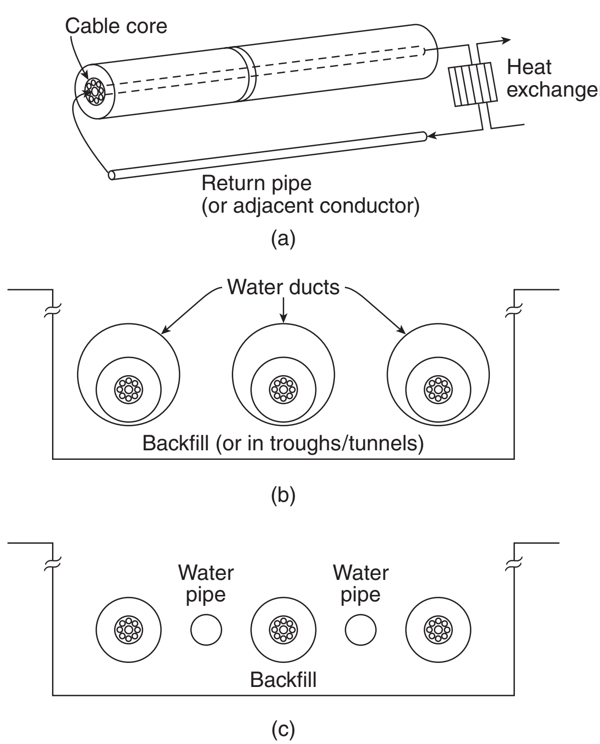


FIGURE 14-83 Alternate methods of cooling single-core cables: (a) internal (core) cooling (SCFF only); (b) integral cooling; (c) lateral cooling.

14.2.10 Cable Uprating and Dynamic Ratings

Cable uprating and upgrading are techniques that are applied to existing cable circuits to get more power through existing infrastructure without building new lines. Many of the basic strategies used for uprating are important to consider for basic ampacity when a new circuit is installed. Utilities frequently need to "uprate" existing circuits; an uprating project can often be implemented quickly to achieve a 10% increase in rating while deferring a multimillion dollar new installation project for several years. In addition, rights-of-way are sometimes difficult to obtain even for underground circuits so getting more capacity through existing lines becomes critically important. Underground cables are especially amenable to uprating, principally because of the very large thermal time constant of the cable/earth system. Also, many older circuits were designed and rated using conservative assumptions; these assumptions may be reviewed carefully, often revealing additional circuit capacity.

Utilities have successfully implemented many approaches for uprating.[20] Typical uprating approaches are summarized below:

- *Characterize the thermal circuit.* In most cases, design engineers properly make conservative assumptions when rating the circuit. A careful audit of route as-built drawings, soil and backfill thermal resistivity, load shape, and so on will provide accurate data and can often permit rerating the circuit 5–10% higher.
- *Perform thermal analysis.* Thermocouples or fiber-optic distributed temperature monitoring cables (Fig. 14-84) can tell the actual cable or pipe temperature at discrete locations or continually along the route. Detailed analysis of loading for the preceding weeks, cable construction data, and trench cross-section information can permit refining the ampacity model, often resulting in significant ampacity increases.[21]
- *Mitigate hot spots.* The thermal analysis may indicate hot spots—a few meters of cable that limit the overall circuit rating—along the route. If the hot spots are localized, they can perhaps be mitigated by

FIGURE 14-84 Distributed temperature system and fiber being used in the field to measure cable temperatures.

removing excess overburden, replacing poor soil with good backfill, recompacting the soil or injecting an additive to reduce thermal resistivity. Sometimes transmission circuits can be upgraded by rerouting nearby distribution circuits or inserting heat pipes to cancel mutual heating effects. Dielectric fluid circulation in HPFF cables can also mitigate localized hot spots.[22] Pipe circuits with only one pipe (no parallel circuit or separate fluid return pipe) can also have some modest circulation using the large fluid reservoir tanks to oscillate fluid in the pipe, mitigating hot spots.

- *Take advantage of load shapes.* Because of its long thermal time constant, the temperature of a cable system today is a result of its loading history for the last several weeks. Utilities typically choose a conservative daily load factor when performing ampacity calculations. Careful calculation of the effects of the actual hourly loads can result in increased ampacities.

- *Fill ducts with a slurry.* The static air space in an XD or SCFF cable duct reduces ampacity by 5–6% or more as compared to a similar direct buried cable. Filling the duct with a removable slurry such as a sand-bentonite mixture can restore that ampacity. Duct ends must be sealed to prevent drying, and it still may be necessary to water jet the slurry if the cable must be removed later. Care must be taken to avoid air pockets, especially if there are dips in the route profile.

- *Reconductor.* Technically called "upgrading," substantial ampacity increases can be achieved for pipe-type cables or XD/SCFF cables in duct by removing the present cable and installing a larger conductor cable, possibly using lower loss insulation (e.g., PPP instead of conventional kraft paper). Utilities sometimes consider allowing a thinner insulation wall on the replacement cable to allow additional space in conduits or pipes for a larger conductor size. Cables with a higher voltage might also be installed, providing a power transfer increase directly proportional with the voltage increase, but this is seldom economical unless the higher voltage already exists in the terminal substations.

- *Add forced cooling.* For HPFF cables, it is possible to remove the liquid from the cable pipe, pass it through a heat exchanger, and send it back into the pipe through a separate liquid line. Although installation and operating costs are high, forced cooling can add 40% to the capacity of an existing cable system. Forced cooling could also be used with XD/SCFF cables using parallel water pipes, but this is uncommon.

- *Dynamic rating.* Cables almost always operate far below their thermal limits, and their conductor temperatures are much lower than allowable values. Therefore, the cables are capable of operating above their steady-state rating for significant periods without exceeding allowable temperatures. In many cases, the cables can even operate above traditional maximum temperatures without unduly decreasing cable life. Dynamic rating systems make this uprating methodology possible by determining the conductor temperature in real time based on measured parameters such as load, ambient earth temperature, and other conditions. A computer can record recent loading data, pipe, shield, or duct temperatures (and fluid temperature for an HPFF system), and ambient earth temperature, and calculates allowable ratings for future periods. The operator has the ability to determine the allowable loading for a certain period of time, determine the length of time permitted before temperature limits are exceeded, and so on. Dynamic ratings can permit ampacity increases of 20%–30%, depending upon operating conditions, with a relatively small investment in equipment.

14.2.11 Soil Thermal Properties and Controlled Backfill

The earth portion of the cable thermal circuit accounts for the greatest percentage of thermal resistance for buried cables—often more than half the total resistance. Native soil thermal resistivity can vary by an order of magnitude along a cable route, and it can vary by a factor of 3 at one location as a function of seasonal moisture content. Accurately characterizing thermal resistivities, developing dry out curves, and choosing the proper value to use in ampacity calculations are important parts of cable system design.

Soil testing is often done in the field (in situ) or laboratory (using samples) to determine the thermal characteristics of the soil in which the cables will be installed. A thermal property analyzer (TPA) is used for these types of measurements (Fig. 14-85). For a thermal resistance test, a thermal needle

FIGURE 14-85 Thermal property analyzer or "TPA." (Courtesy of Geotherm, Inc.)

consisting of a heater and thermister is placed into the soil sample (Fig. 14-86). During the test, a constant heat output is applied to the thermal needle while the change in temperature is recorded; the slope of the recorded time-temperature curve is proportional to the soil thermal resistivity. This information is then used for ampacity calculations and to design the backfill materials to place in the cable trench.

In the laboratory, thermal dry-out curves may also be prepared that shows the effect of soil moisture content on thermal resistivity. Generally, as soils have increased moisture content, the thermal resistivity decreases (Fig. 14-87). Since cables produce heat, some drying should be considered when evaluating ampacities.

Significant ampacity increases can be achieved at reasonable cost by placing a controlled backfill with a low, stable thermal resistivity around the cables, ducts, or pipes. A good controlled backfill is characterized by low-thermal resistivity, less than 1 K°-m/W, when completely dry, and a fairly flat curve of thermal resistivity versus moisture content. The proper material should be designed for each project using locally available materials, when possible, to provide the best installation and thermal characteristics at a reasonable cost. Well-graded sands or limestone screenings have been used since 1950s. Recently, many utilities have been installing FTB, which is a low thermal resistivity, free-flowing engineered material consisting of natural mineral aggregates, sands, cement, water, and a fluidizer (Fig. 14-88). It is delivered in ready-mix concrete trucks, and flows throughout the trench so there is no need for compaction.

14.2.12 Electrical Characteristics

Calculation of Electrical Parameters. Electrical characteristics for cables are calculated using various established methods. Resistance, capacitance, and inductance can be calculated using the industry

FIGURE 14-86 Thermal probe being inserted into an auger during geotechnical testing.

standard procedures for cable ampacity.[7] Sequence impedances for cables can be approximated using various sources.[23,24] Rigorous calculations for pipe cables are difficult because of the presence of steel (ferromagnetic) pipe; Ref. 24 has traditionally been considered the best approach available for pipe type, though it is known to be in error in many cases because of the great variability in line pipe permeability. For a typical single-core (XD, SCFF) transmission cable installations, a detailed procedure exists.[25] but does not easily lend itself to hand calculations. Induced sheath voltages may be calculated using the appropriate IEEE standard.[4]

Although surge impedance loading is never an issue for cables (unlike overhead lines) because cables ratings are always constrained by thermal limits, system studies sometimes require the surge impedance for modeling power systems. This can be calculated using the following formula where f is the power frequency, ε is the dielectric constant of the insulation, D_i is the diameter of the cable over the insulation, and D_c is the diameter of the cable conductor.

$$Z_S = \frac{f}{\sqrt{\varepsilon}}\, ln\left(\frac{D_i}{D_c}\right) \qquad (14\text{-}99)$$

Fault current capability of a cable can be an issue, particularly in stiff power systems. The fault current capability of underground cables can be calculated assuming adiabatic conditions using the

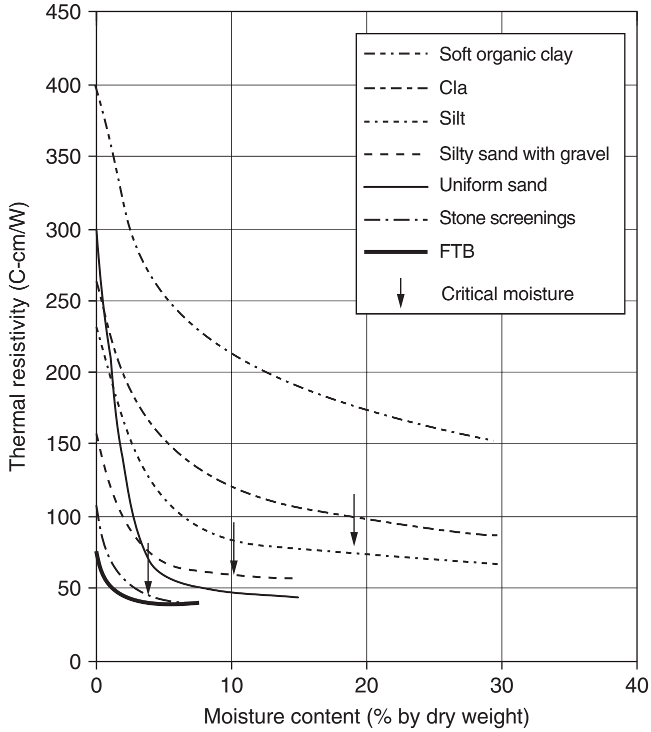


FIGURE 14-87 Soil thermal dry-out curves.

FIGURE 14-88 FTB being installed in a cable trench.

following equation where conductor area is in square millimeters, and time is in seconds (i.e., at 60 Hz, 30 cycle clearing time will be 0.5 s). T_1 is the temperature of the sheath or conductor prior to the fault in °C, T and k are material parameters for conductor or sheath material (see Table), and T_2 is the allowable maximum temperature for the insulation or jacket material (whichever is lower) in contact with the conductor or sheath. Conductor cross-sectional area is specified in square millimeters.

$$I_{SC} = (\text{Conductor Area})\sqrt{\frac{k}{time}ln\left(\frac{T_2 - T}{T_1 - T}\right)} \tag{14-100}$$

Conductor/Sheath Material	Constant, k	Inferred Temperature of Zero Resistance, T °C
Copper	50237	−234.5
Aluminum	21144	−228.0
Lead	1641	−236.5

Though polyethylene jackets are often used, many manufacturers allow higher shield/sheath temperatures than 150°C (some up to 250°C) based on experimental evaluation and the conservative adiabatic assumptions.

Insulation/Jacket Material	Temperature, T_2 °C
High-density polyethylene	150
Cross-linked polyethylen	250
Polyvinyl chloride	150
Ethylene propylene rubber	250
Impregnated paper	150
Laminated paper-polypropylene	150

14.2.13 Magnetic Fields

Magnetic fields are generated by underground cables as a result of current flowing in the conductors of the cables. The intensity of magnetic fields is a function of conductor and shield currents, phase spacing, and distance from the source and can be calculated using the Biot-Savart Law as follows:

$$B = \frac{\mu_0 I}{4\pi}\int \frac{dS \times r}{r^2} = \frac{\mu_0 I}{2\pi r} \tag{14-101}$$

with μ_0 is the permeability of free space ($4\pi \times 10^{-7}$ Wb/m^2), I is the conductor current in amperes, and r is the distance from the conductor in meters. Magnetic fields are reported in micro-tesla (μT) or milli-Gauss (mG); 1 T is 10^4 G, so 1 mG is equal to 0.1 μT.

Magnetic fields generated by 3-phase cable circuits must account for the individual cable phase current magnitude and phase angle and that the magnetic field is actually the resultant of the minor and major (real and imaginary) axes of the rotating elliptical magnetic field phasor. This can be done easily with the use of a computer for cables that do not have ferromagnetic elements such as the steel armor on submarine cables, the metal casing often used with directionally drilled cables, or the steel pipe around pipe-type cables. Cable systems with ferromagnetic components should be modeled using finite element software to account for the nonlinear nature of the problems and the variability in the permeability of steel with field intensity. CIGRE has developed procedures to calculate magnetic fields for cables with[26] (HPFF) and without[27] (XD, SCFF) ferromagnetic components; these methods use empirical relationships to account for the shielding effects of the iron-based shielding.

The importance of magnetic field management is a topic that varies depending largely on public perception of the issue. Some utilities have investigated various magnetic field management methods to find the best solution for a particular application.[28]

14.2.14 Installation

XD and SCLF Cables. Installation of single-core transmission cables (SCLF and XD) requires that the cables either be directly buried or installed in conduits, although they are sometimes installed in troughs or mounted on the inside of tunnels. Preferably, the cables are installed direct buried—usually to a nominal depth of 1 m cover over the top cable—with center-line spacing of approximately 300 mm. Joint bay locations for direct buried cables or manholes for conduit systems are dictated by cable reel lengths (typically limited to 2600 ft or 800 m for large cables), local restrictions on the permissible length of open trench, or by maximum permitted sheath voltages. Depending upon the route plan and profile, allowable pulling tensions or sidewall pressures can limit spacing for cables installed in conduit. When cables are to be installed in conduits or pipes, it is important to consider the maximum pulling tension that may be encountered during the installation.[29] Fluid reservoirs for SCLF cables are spaced according to elevation changes or maximum transient hydraulic pressures during load cycling. Generally, each substation end has a set of reservoirs, although they may be placed in manholes as needed. Reference 30 gives details of a typical SCLF cable installation.

For conduit systems, the ducts are installed first and the cables pulled in and splicing done later (Fig. 14-89). This type of installation has benefits for urban environments where lengths of open trench must be limited and surface restoration must be done promptly. Many utilities install spare conduits so that additional cables or other utilities can be installed along the same conduit run.

FIGURE 14-89 Duct bank being backfilled.

FIGURE 14-90 Cable trench with final FTB layer installed.

Direct buried installations require that the trench be opened for an entire pull section. Depending on cable dimensions and reel limitations, this can be 500–800 m of trench. The trench frequently must have sheathing and bracing installed to prevent itself from collapsing. Where thermal sand or FTB will be used, a layer of the material is placed in the trench prior to placing cable rollers along the route. Then, the cable is pulled from one end to the other along the rollers using a winch. Once all cable phases have been pulled, thermal sand or FTB is placed around and above the cables (see Fig. 14-90). Additional backfill, a concrete cap, or marking tape can then be placed on top of the envelope around the cables before restoring surface conditions.

14.2.15 HPFF Cables

Pipe-type cables require much more complex installation procedures and specialized equipment than extruded-dielectric cables.[23, 31] They are most commonly installed in city streets, where the narrow trench requirement, speed with which the pipe sections (approximately 12 m) can be installed, and ruggedness of the pipe offer advantages. Trench openings more than 200 m are preferred, and pipe installation can progress at several hundred feet per day, depending upon subsurface congestion, traffic conditions, and so on. The trench can be backfilled (often with a controlled backfill) as soon as a pipe section is placed, welded, the weld tested, coated, and the integrity of the pipe coating is checked. Manhole-to-manhole sections of 700–1000 m in length can be installed, vacuum/pressure tested, and pressurized with dry nitrogen until cable is installed later.

Cable is delivered to the site on large sealed steel reels and placed in a special trailer as shown in Fig. 14-91, lagging is removed from the reels, the three phases are brought together into a single pulling yoke, and the cable is pulled to the adjacent manhole. Characterizing the pulling tension is important to avoid damage to the cable or jamming of the cables in the pipe. One critical factor to consider is the coefficient of friction to use for the cable pipe and pulling rope.[32] Pulling cables through the steel pipe is accomplished using a specialized pulling winch (see Fig. 14-92). Pressure-tight "night caps" are placed over the cable ends, the pipe is evacuated, and slightly pressurized with a nitrogen gas until splicing begins. A splicing trailer is placed over the manhole to provide a work area and maintain humidity control in the manhole. A 345-kV splice takes 5–7 days to complete.

Each termination takes about a day to assemble. Humidity-controlled enclosures are used for 345- and 230-kV cables, and occasionally for 138-kV cables. For both splices and terminations, it is important to maintain a dust-free and low-moisture environment during installation work.

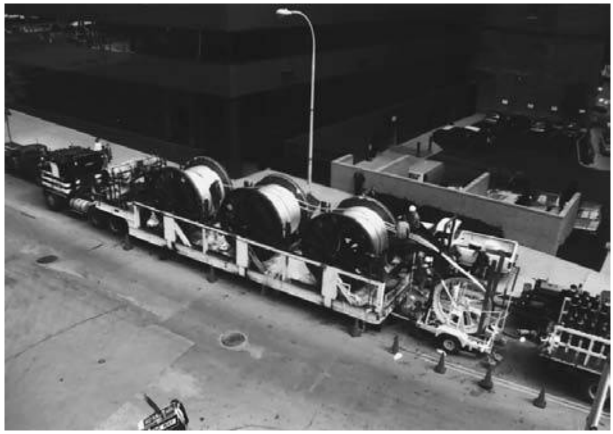


FIGURE 14-91 Reel trailer.

Pressurizing plants for HPLF cables are generally assembled in weatherproof enclosures at the manufacturing facility, shipped to the site and set in place on foundations (which may include a moat), where electrical and hydraulic connections are made. Before they are energized, the cables must be pressurized very slowly to prevent insulation damage and ensure that any minor amounts of gas in the cable are dissolved in the dielectric liquid.

14.2.16 GITL

Since the enclosures are rigid and diameters are large (e.g., 40 cm for a single phase of a 345-kV system), this cable type does not have as much flexibility as extruded-dielectric or pipe-type cables. Bends of more than a few degrees require a specially made elbow. Many systems have been installed directly buried, but most new designs call for installation in a trough that minimizes corrosion concerns and allows rapid access for maintenance. Care must be taken to properly design for thermal expansion of both the enclosure and the conductor.

FIGURE 14-92 Pulling winch trailer.

FIGURE 14-93 GITL installation. (Courtesy of CGIT/AZZ.)

The compressed-gas-insulated system is seldom installed in city streets; most applications are within substations or power plants, or on the utility right-of-way. Figure 14-93 shows a typical installation.

14.2.17 Special Considerations

Submarine Installations and Water Crossings. Short water crossings, less than 2 km, can be accomplished by horizontal directional drilling in many cases, and the cable system can be pulled into the casing pipe or directly into the bore in some instances. Traditional cut-and-cover installations are suitable where water bottom conditions and environmental considerations permit. Extruded-dielectric, pipe-type, and self-contained cables can be suitable for these short crossings.

Longer water crossings have sometimes been installed by laying the cables on the bottom, but the high incidence of ship anchor and trawler sled damage have caused almost all new installations to be buried by trenching, jetting, or plowing using special equipment to embed the cable below the water bottom. The cable laying ship, with special navigation and positioning equipment, pays off cable from a large turntable to an embedder that augers or water jets a trench for the cable to fall into during the laying operation (Fig. 14-94). The cable laying ship has the capacity to hold several miles of cable, depending on the size, which minimizes the number of splices that must be installed.

Thermal analysis of water-bottom material is important, especially for lake and river crossings where sediments tend to accumulate. Cables have failed because of drying out of sediment material around the cables, even though the water depth at the failure location was more than 10 m.

Self-contained liquid-filled cables have been used for major ac installations.[33, 34] Extruded-dielectric cables are becoming more common for ac submarine cable installations.[35] Solid-type cables are commonly applied for long dc cable installations.[36] The 1990s and 2000s have seen a great deal of interest in submarine cable applications, worldwide.

Horizontal Directional Drilling (HDD). Horizontal directional drilling is becoming a common way to install cables under rivers, major highways, and other locations where trenching is not feasible.

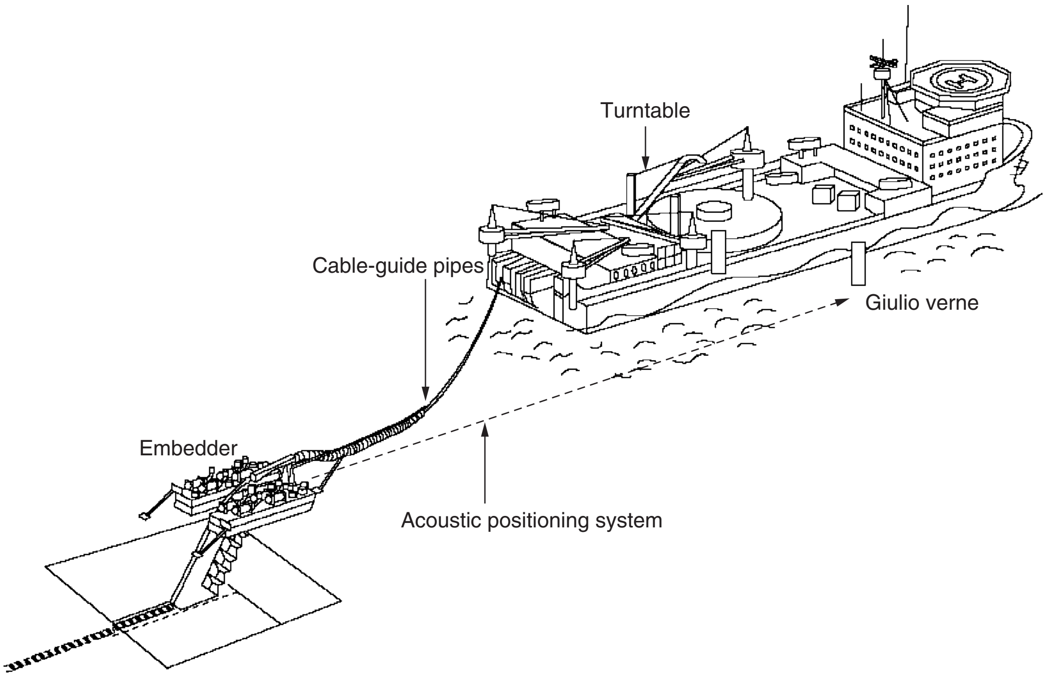


FIGURE 14-94 Cable laying ship and sea bottom embedder. (Courtesy of Prysmian.)

Drilling unit capabilities have increased, and prices have decreased, as the technology has matured. A recent installation in South Carolina successfully installed a pipe-type cable along a 7100 ft (2.2 km) HDD crossing. HDD installations are even being made in city streets, to avoid other utilities, to minimize traffic disruption, and to minimize the amount of contaminated soil that may be encountered.

Careful design work, including characterization of the materials along the drill path, must be performed so the proper drilling equipment can be selected to ensure a successful installation without loss of drilling fluid to the surface. Detailed thermal analysis is required to properly account for the 10–30 m burial depth. The utility must decide whether to install a casing pipe, either plastic or steel, which will require a larger bore and may decrease ampacity but will help ensure successful cable installation. Entry and exit angle requirements of 8–12 degrees can require a large setback for deep borings. Figure 14-95 shows a guided boring rig for a river crossing. Staging areas can be large, especially since casing pipe and ducts or cable pipe should be madeup beforehand and installed without stopping—the drilling mud may begin to set up.

Since the cable will be inaccessible for repairs, utilities should consider stocking a section of spare cable that is the full length of the bore. For pipe-type cables, or three XD cables installed in a single duct or pipe, the spare cable must be 3 times the length of the bore. In the case of extruded cables, some utilities have installed a fourth cable in a spare duct, to speed reenergization in the event of cable failure.

14.2.18 Accessories

General. Although accessory cost is usually a small percentage of the total project cost, proper selection and application of accessories are extremely important for long-term reliability of underground cable systems. Historically, outages due to accessories are more frequent than those due to

FIGURE 14-95 Guided boring equipment for a river crossing.

problems in the cable itself. Much of the development activity, especially in extruded-dielectric cables, has been directed at simpler, easier-to-install, reliable splices, and terminations. (Note that a splice is the electrical reconstructing of the electrical insulation and shielding. A joint also includes housing, fluid supply, and so on.)

XD Cables. Different splice designs have been applied as the industry has matured, including taped splices, taped and molded splices, prefabricated and premolded splices, back-to-back SF_6 potheads, and a miniature extrusion facility set up in a manhole. In the late 1990s, premolded joints became the standard for field installation at most voltage levels. These joints were fully tested in the factory, and they required less assembly time in the field. In cross-bonded systems, the splice must accommodate an insulator in the outer shielding, to isolate electrically the shields/sheaths of adjacent cable sections.

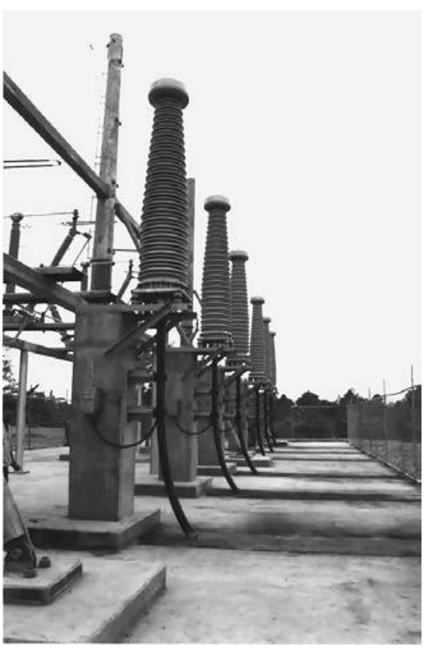


FIGURE 14-96 Terminations mounted on concrete risers.

Terminations are typically slip-on, with a dielectric liquid to fill any voids in the high-stress regions. The area above the liquid contains a compressible sponge, or simply an air space, to accommodate fluid expansion. Some designs require a small fluid reservoir for this purpose. Porcelain housings have traditionally been used, but polymer housings are becoming more common, especially for applications where the terminations are mounted directly on riser poles as shown in Fig. 14-96. Terminations are available for air installation, installation in single-phase SF_6 bus, and installation in 3-phase SF_6 bus.

Link boxes permit connections among the sheaths for grounding and cross bonding. The links must be removed when the cable jacket is tested electrically as part of routine maintenance. A sheath voltage limiter, resembling a distribution-class surge arrester, is often placed at the nongrounded end of a cable sheath to limit overvoltages, which might damage the jacket.

HPFF Cable Splice. These types of cable splices are normally individual hand-taped assemblies which are bound together using an aluminum "spider" to maintain spacing and reduce flexing. They are contained in a 3- or 5-piece welded steel joint casing that can be 40 cm diameter and 3.5 m long. Trifurcating splices provide the transition from three cables in a steel pipe to individual cables in stainless-steel pipes rising up to a termination. Special "stop joints" are installed on high-pressure liquid-filled (HPLF) cables to minimize fluid loss in the event of a major leak, and to permit maintenance on the cable without draining large amounts of liquid.

Terminations (also called potheads) have paper "piano rolls" to help grade the electrical stress. Many 230-kV terminations and all 345-kV terminations have capacitor stacks inside the termination body to smoothly grade the electrical stress from ground to line potential. The HPLF terminations are pressurized with liquid from the cable pipe through filters and check valves that limit liquid loss in event of termination failure. HPGF terminations are typically just pressurized with nitrogen gas from the cable pipe, although some designs have a liquid-filled termination.

Pressurizing plants are sized to accommodate the maximum liquid demand and expulsion under all loading conditions. The plants contain a reservoir tank (typically 8000–40,000 L), pressurizing pumps, controls, and alarms. Since pressurization is critical to keep a line in service, backup pressurizing pumps are provided, sometimes including nitrogen-driven pumps that can operate during power outages. Emergency generators are supplied in some instances.

Cathodic protection is applied to protect the steel pipe in case the corrosion coating is damaged. The system consists of a rectifier to provide the -0.85 V needed to ensure that current flows from the earth to the pipe, plus a polarization cell (or solid-state equivalent) to provide a low-impedance path to earth for fault currents.

SCLF Cables. Normal joints are single-phase assemblies that have a special connector or ferrule to permit liquid flow through the conductor core. They are spliced in the same manner as a pipe-type splice and have a nonmagnetic lead, copper, or aluminum housing. Insulators are placed in the housing to permit cross bonding of the cable sheaths. Stop joints are placed every 1200–3000 m to isolate the liquid-supply sections. These joints may have to be placed closer together in areas with substantial elevation changes.

Terminations are very similar to those for pipe-type cables. Trifurcating joints are generally not required since almost all transmission-voltage self-contained cables are single-conductor design.

Liquid reservoirs are spaced along the circuit. Distances depend upon elevation changes, cable core diameter, and maximum allowable pressures. Typical spacings are 1500–5000 m. These reservoirs have alarms to indicate low liquid level or pressure. These alarms must be run back to the substation, typically in a separate communication duct, to notify operators of cable system problems.

GITL accessory requirements. These accessory requirements are minimal. Splices are an integral part of the assembly, and terminations are factory-made and attach to the cable in a similar manner as to a normal joint. The system contains alarms to warn of low pressure or high moisture content. Joints occur every 12–18 m (see Sec. 14.2.15) but are simple and straightforward. Expansion and contraction of the insulating gas is accommodated by allowing the pipe pressure to increase or decrease. Thermal-mechanical expansion of the aluminum bus is permitted by the use of bellows at bus connection points. The aluminum enclosure must be protected from corrosion using protective coatings and cathodic protection if the system is buried.

14.2.19 Manufacturing

Manufacturing of transmission cable should be done using quality materials and an environment free of contaminants. Impregnated-paper insulation technology has existed since the late 1800s and is the basis for substantial use of HPFF transmission cable in the United States. Extruded dielectrics, particularly cross-linked polyethylene (XLPE) insulation, have been used increasingly in recent years as the manufacturing technology to use this type of insulation up to 500 kV is developed.

Manufacturing requires that copper or aluminum wire be drawn and wrapped together for the required conductor cross section. Depending on the size of the conductor, the conductor may be formed in segments—typically four for HPFF and XD cables and six for SCLF cables—that must be laid together. A binder tape may be applied to hold the segments together, and a taped shield is applied over the conductor on paper cables.

Insulation is applied over the conductor by drawing the conductor through an appropriate type of machine. For paper-insulated cables, the paper tapes are applied dry by wrapping the insulation around the conductor until sufficient insulation has been built up for the operating voltage. The paper cable is then heated under vacuum to remove moisture before dielectric liquid is applied in an impregnation tank. XD cables have insulation applied using a catenary, horizontal or vertical extrusion machine. A "triple extrusion head" is a modern approach to applying extruded insulation where the conductor screen (shield), insulation, and insulation screen (shield) are applied in one process.

After the insulation is applied and impregnated, HPFF cables are often wrapped with a Mylar tape intercalated with a metallic shielding tape. Then skid wires are applied helically over the tapes to prevent mechanical damage to the cables when the three phases are pulled into the pipe. The cable is placed onto shipping reels, and the reels are sealed and blanketed with dry nitrogen to limit gas and moisture ingress.

In single-core cables, bedding may be placed over the insulation screen to prevent damage to the cable during thermal expansion of the cable. If a wire shield is to be installed, this is normally put in direct contact with the outer insulation shield prior to applying the bedding layer. Additional bedding in SCLF cables or a water-swellable tape or powder in XD cables is then applied before a metallic sheath. The sheath is most commonly made of lead, corrugated stainless steel, copper, or aluminum, and can be applied in an extrusion or using a seam weld and a press. Finally, a jacket made of polyethylene or polyvinyl chloride is extruded over the sheath. The cable is placed on a cable reel with the cable ends sealed to prevent moisture from getting into the cable. Typically, a pulling eye is attached at the factory for use during installation.

14.2.20 Operation and Maintenance

Operation and maintenance requirements of underground transmission cables are small compared with most electrical power equipment and are essentially confined to the dielectric fluid-handling systems for HPFF and SCLF cables, the outer cable jacket or pipe protective coating, and the corrosion protection systems for HPFF cables.

Extruded dielectrics and SCLF transmission cables have an outer jacket typically constructed of polyethylene or polyvinyl chloride. The jacket is periodically tested to verify that there is no damage. Damage to the jacket may result in localized corrosion and unexpected circulating currents in the sheath that will cause I^2R losses in the sheath, lowering the effective ampacity or overheating the cable. If the cables are cross-bonded, the link boxes should be checked to be sure that the sheath voltage limiters are not damaged and there is good ground continuity for the local earthing (Fig. 14-97).

All SCLF and HPFF terminations and some XD terminations contain small amounts of dielectric liquid (200–600 L). The terminations should be inspected to be sure that there are no fluid leaks. In areas with high levels of air pollution, such as near the ocean, near unpaved roads, or in industrial areas, it is important that the outside surface of the terminations be cleaned to avoid surface tracking. Porcelain terminations should be free of cracks or other mechanical damage.

Dissolved gas analysis is a useful tool for HPFF and SCLF cables to determine the condition of a cable by examining the ratio of certain gases in the dielectric liquid. Field dissipation factor measurement can also be performed.[37] For both dissolved gas analysis and dissipation factor measurements, the trend with time rather than an absolute value is of interest for monitoring any degradation with time. Many utilities perform tests once per annum. Joints in locations susceptible to thermomechanical bending should be x-rayed every few years. Semiannual or quarterly dew-point readings should be taken on CGIT lines.

A few utilities have experienced problems with dielectric fluid leaks in HPFF cables caused by a combination of several factors including coating failure, cathodic protection problems, and stray currents in the earth. Monitoring of cathodic protection system performance and periodic testing of the pipe coating can limit the possibility of leaks, and leak detection and location using per-fluorocarbons tracers or leak detection wire buried along the cable pipe have been used with success.

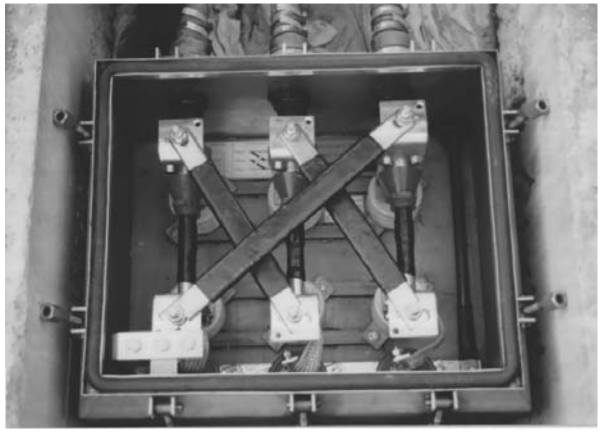


FIGURE 14-97 Cross-bonding box with links installed.

14.2.21 Fault Location

Although they are very uncommon, cable failures can be difficult to locate as compared to overhead lines where it is possible to visually inspect the entire line route. Locating cable faults requires the application of two types of techniques: terminal methods and tracer methods. Terminal methods are applied from one or the other end of the faulted cable and give an approximate location of a failure. Terminal methods include cable radar, bridge techniques, and time domain reflectometry. Tracer methods involve putting some type of signal on the faulted cable and then detecting changes in the signal when in close proximity to the fault location. Tracer techniques include tone sets, earth gradient, capacitive discharge ("thumping"), and magnetic pulse detection. Many times, the location of a failure is approximated with a terminal technique and then followed by a tracer technique precisely to locate the failure.

Pipe-type cable failures are particularly challenging. If fluid circulation is used on an HPFF cable, circulation should be stopped immediately once there is an indication of a failure to avoid contaminating the dielectric fluid. Since visible damage will generally not be apparent once the detection method has located the apparent fault location, the cable pipe must be x-rayed to determine if the exact location of the cable failure has been found. The liquid in the cable pipe must then be frozen on either side of the failure before the pipe can be cut and a repair initiated.

A good summary of classical fault location techniques can be found in Ref. 38. The IEEE Insulated Conductors Committee Working Group 12-48 is developing a guide for fault location.

14.2.22 Corrosion

Corrosion of underground cables can result from two phenomena; galvanic and electrolytic corrosion. Galvanic corrosion results from the electrochemical interaction of two or more materials, such as zinc and copper, with the level of galvanic action depending on the galvanic potential of the two materials. Steel, such as in a HPFF cable, can undergo corrosion from oxygen in acidic soil. Galvanic corrosion can be prevented by isolating the cable or pipe from other materials by using a protective coating or jacket. Pipe-type cables are often connected to a cathodic protection system where a negative voltage (relative to ground) is applied to the pipe using a rectifier and a polarization cell (or its solid-state equivalent is used) to provide a low-impedance path to ground for fault currents. Sacrificial anodes can also provide protection in the event the pipe-coating is damaged. Some utilities actively monitor the cathodic protection systems on the cable pipe to detect damage to the pipe coating as early as possible.

Electrolytic corrosion results from an externally imposed current leaving a conductive element where an anode zone is created. This might occur on the armor of a submarine cable. Protecting a cable from an electrolytic corrosion can be difficult because it is generally not possible to eliminate the source of the stray current that causes electrolytic corrosion. Instead, imposed cathodic current protection (ICCP) is used. In this scheme, anode beds are connected to the positive terminal of a rectifier, and the cable armor, for example, is connected to the negative terminal. Anode beds are placed along the cable to produce an imposed current on the cable armor. The level of stray currents will designate the current density, which must be produced by the ICCP system and anode beds in order to protect the cable.[39]

14.2.23 Testing

There are various tests performed on underground cable systems during design, after manufacture, before and after installation (commissioning tests), and as part of maintenance. Various IEEE, IEC, AEIC, and ICEA standards have been developed that address this subject.

A manufacturer often performs a long-term "type test" prior to commercial production to prove the integrity of a cable and its accessories. Some purchasers of a cable system may also require that a type test be performed as part of their cable order, generally when a large cable system will be supplied or when a new set of accessories is being used with a cable for the first time.

As part of manufacturing, routine and sample tests are performed and verified by the purchaser's representative to confirm that the cable was manufactured according to the designated specifications. Accessories also undergo routine tests and some purchasers choose to observe these tests or at least review completed test reports. Routine tests may include partial discharge (PD) testing, particularly on extruded cable systems, as well as capacitance, conductor resistance, dissipation factor (tan δ), and jacket integrity tests on extruded cables. Pipe-type cables on shipping reels may have the dew point and nitrogen pressure tested prior to shipping, and self-contained cables may have a check of dielectric fluid pressure rise to be sure that there is no air in the system. Some of these tests may be performed just prior to cable installation after cable reels are on site to be sure that the cable reels were not damaged during shipping.

After laying tests (commissioning tests) may include checks of conductor resistance, cable phasing, jacket integrity tests (on extruded or self-contained cables), a dc high-potential ("hi-pot") test on paper insulated cables (only), as well as tests on dielectric fluids before filling cable pipes or pressure reservoirs. XLPE-insulated cables generally are subjected to an ac "soak" test (energized without carrying load) or may be subjected to some higher voltage if available within the connected substation. Very low-frequency (VLF) tests sets may be employed for testing XLPE cables although they have limited availability at the higher transmission voltages. Utilities are increasingly using variable frequency resonant test sets to apply multiples of rated voltage to an XLPE cable system where the higher voltage is not available locally; this has some advantages including very low-fault current energy to apply to a fault in the event a problem is detected during the test.

Tests during operation and maintenance include periodic jacket integrity tests, PD, dissipation factor measurements (mainly on paper-insulated cables), and the various methods used to locate cable faults as described earlier. Dielectric fluid testing is performed on paper cable to check for dissolved gases (dissolved gas analysis or "DGA"), which is often a good indicator of various problems with a cable system.

14.2.24 Future Developments

Research & Development (R&D) Efforts. R&D work on cable systems themselves has waned as the industry entered the twenty-first century. As some paper-insulated transmission cables reached 40 years of operation, studies were done in the mid 1990s to evaluate loss of life and end of life effects. Many cable manufacturers continue to push the upper voltage limits for cross-linked polyethylene cable systems with systems available up to 500 kV from several suppliers. Superconducting cables, which had been evaluated and set-aside in the 1970s, have moved past the prototype stage and are seeing limited commercial applications. Manufacturers are pursuing more practical ways of offering HVDC cable systems, but most of this work relates to the connected converter stations rather than the cables themselves; one manufacturer offers a polymeric dc cable for this purpose.

Magnetic fields were once a critical issue for many utilities, but technical interest in this topic no longer seems to be a high priority. However, dynamic ratings and uprating methods are of great interest since these methods generally allow utilities to run their cable systems closer to the ultimate limit.

REFERENCES

1. Bascom, E. C. III, D. A. Douglass, G. C. Thomann, and T. Aabo, Hybrid Transmission: Aggressive Use of Underground Cable Sections with Overhead Lines, CIGRE, Paris, France, 1996.
2. Stewart, J. A., J. A., Williams, D. D. Wilson, System Implications of Underground Cable Systems, IEEE Transmission & Distribution Conference, Anaheim, CA, September 15–19, 1986.
3. Williams, J. A., Overhead versus Underground Analysis, Proceedings, Transmission & Distribution World Expo 97, Atlanta, November 11–13, 1997.
4. *IEEE Guide for the Application of Sheath-Bonding Methods for Single-Conductor Cables and the Calculation of Induced Voltages and Currents in Cable Sheaths*, ANSI/IEEE Standard 575-1988 (under revision).
5. Auerbach, R. H., et al., 525-kV Self-contained Oil Filled Cable Systems for Grand Coulee Third Power Plant—Design, Test, Manufacture, and Installation, IEEE Transactions on Power Apparatus and Systems, vol. 99, no. 5, September/October 1980, pp. 1722–1733.
6. Graybill, H. W., J. A. Williams, Underground Power Transmission with Isolated-phase Gas-insulated Cables, *IEEE Transactions on Power Apparatus and Systems*, vol. 89, no. 1, January, 1970. pp 17–23.
7. *Calculation of the Continuous Current Rating of Cables (100% Load Factor)*, International Electrotechnical Commission, publication 60287, 1994.
8. Neher, J. H., M. H. McGrath, The Calculation of Temperature Rise and Load Capability of Cable Systems, Transactions of AIEE, Power Apparatus Systems, vol. 76, October 1957, pp. 752–772.
9. Neher, J. H., A simplified Mathematical Procedure for Determining the Transient Tempeature Rise of Cable Systems, *Transactions of AIEE*, Power Apparatus Systems, August 1953, p. 712.
10. *Calculation of the Cyclic and Emergency Current Ratings of Cables*, International Electrotechnical Commission, 1st ed., publication 853–2, 1989.
11. IEEE Standard Power Cable Ampacity Tables, IEEE Standard 835-1994.
12. *Electra*, Calculation of the Continuous Rating of Single-Core, Rigid-Type, Compressed-Gas-Insulated Cables on Still Air with No Solar Radiation, *Electra*, No. 100, 1985, pp. 65–75.
13. *Electra*, Calculation of the Continuous Rating of Single-Core, Rigid-Type, Compressed-Gas-Insulated Cables on Still Air with No Solar Radiation, *Electra*, No. 106, 1985, pp. 23–31.
14. House, H. E., P. D. Tuttle, Current-Carrying Capacity of ACSR, *AIEE Transmission & Distribution Committee*, Paper 58-41, February 1958, pp. 1169–1177.
15. *IEEE Standard for Calculating the Current-Temperature Relationship of Bare Overhead Conductors, IEEE Standard 738–1993*, November 1993.
16. Neher, J. H., The Transient Temperature Rise of Buried Cable Systems, Paper 63-917, *IEEE Insulated Conductors Committee*, June 16–21, 1963.
17. Buckweitz, M. D., D. B. Pennell, Forced Cooling of UG Lines, *Transmission & Distribution*, vol. 28, no. 4, April 1976, pp. 51–58.
18. *Electra*, Calculation of Continuous Ratings for Forced Cooled Cables, CIGRE WG-08, Study Committee No. 21, *Electra*, no. 66, pp. 59–84.
19. Purnhagen, D. W., Designer's Handbook for Forced-Cooled High-Pressure Oil-Filled Pipe-Type Cable Systems, *EPRI Report No. EL-3624*, July 1984.
20. Bascom, E. C., III, Underground Cable Uprating and Upgrading Tutorial, *Transactions of IEEE PES Transmission & Distribution Conference*, Paper 03TD0362 (Panel Session), Dallas, Texas, 7–12 September 2003.

21. Bascom, E. C., III, J. A. Williams, Taking Your Cable's Temperature, *Transactions of the Transmission & Distribution World Expo*, 7–9 May 2002, Indianapolis, Indiana.
22. Williams, J. A., E. C. Bascom, III, and T. Aabo, Field Test Program and Results to Verify HPFF Cable Rating, *IEEE PES Transmission & Distribution Conference*, September 22–27, 1991.
23. *Underground Transmission Systems Reference Book*, Electric Power Research Institute, TR-101670, 1992.
24. Neher, J. N., The Phase Sequence Impedance of Pipe-Type Cables, *IEEE Transaction Power Apparatus & Systems*, vol. 83, August 1964, pp. 795–804.
25. Lewis, W. A., G. D. Allen, Symmetrical-Component Circuit Constants and Neutral Circulating Currents for Concentric-Neutral Underground Distribution Cables, *IEEE PAS*, vol. PAS-97, no. 1, January/February 1978.
26. *Electra*, Magnetic Field Calculation in Underground Cable Systems with Ferromagnetic Components, Joint Task Force 36.01/21, *Electra*, October 1997, no. 174, CIGRE.
27. *Electra*, Magnetic Fields in HV Cable Systems 1: Without Ferromagnetic Components, Joint Task Force 36.01/21, Electra, June 1997, no. 174, CIGRE.
28. Bascom, E. C., III, J. H. Cooper, W. Banker, R. Piteo, A. M. Regan, and S. A. Boggs, Magnetic Field Management Considerations for Underground Cable Duct Bank, *Transactions of IEEE PES Transmission & Distribution Conference*, Paper 05TD0399, Dallas, Texas, 21–26 May 2006.
29. Rifenburg, R. C., Pipe Line Design for Pipe-Type Feeders, *Transactions of the AIEE*, December, 1953, pp. 1275–1288.
30. Kozak, S., J. T. Corbett, and F. J. Bender, Features of the New 138-kV Self-Contained Oil-Filled Cable System for Detroit Edison, *IEEE Transactions Power Apparatus & Systems*, vol. PAS-94, May/June 1975, pp. 949–958.
31. Hatcher, C. T., R.W. Gillette, and R. W. Burrell, 345-kV Underground Transmission on the Consolidated Edison Company of New York System, *IEEE Transactions on Power Apparatus and Systems*, vol. 85, no. 4, April, 1966, pp 353–360.
32. Bascom, E. C., III, J. R. Stewart, and L. Y. Tang, Evaluation of Pulling Rope Wear and Coefficient of Friction for Pipe-Type Cables, *IEEE Transactions on Power Delivery*, vol. 12, no. 2, April 1997.
33. Foxall, R. G., K. Bjorlow-Larsen, and G. Bazzi, Design, Manufacture and Installation of a 525-kV Alternating Current Submarine Cable Link from Mainland Canada to Vancouver Island, CIGRE, Paper 21-04, Paris. 1984 session.
34. Grzan, J., E. I. Hahn, R. V. Casalaina, and J. O. C. Konsog, The 345-kV Underground and Underwater Long Island Sound Project, *IEEE PAS*, vol. 8, no. 3, July 1993.
35. Cooper, J. H., M. J. Polasek, Planning and Installation of the 138 kV South Padre Island Submarine Cable, *IEEE Transactions on Power Delivery*, vol. 8, no. 4, October 1993.
36. Lawson, W. G., W. Zawawie, B. Lunyong, Z. J. Neag, W. Stangl, and J. J. Vithayathil The Power Transmission System of the Bakun Hydroelectric Project, Paper Presented at the *World Energy Council Seminar*, Kuala Lumpur, Malaysia, December 3–4, 1996.
37. Cooper, J. H., G. A. MacPhail, S. Cherukupalli, et al., Insulation Condition Tests On Two Underground Transmission Cable Circuits at BC Hydro, *Transactions of Doble Engineering Conference*, Spring, 1996.
38. Bascom, E. C., III, D. W. Von Dollen, and H. W. Ng, Computerized Underground Cable Fault Location Expertise, *IEEE Transmission & Distribution Conference*, Chicago, 1994.
39. Bascom, E. C., III, et al., Construction Features and Environmental Factors Influencing Corrosion on a Self-Contained Fluid-Filled Submarine Cable Circuit in Long Island Sound, *IEEE Transactions on Power Delivery*, PE-034-PWRD-0-06, January 1997.

SECTION 15
DIRECT CURRENT POWER TRANSMISSION

Michael P. Bahrman
ABB, Inc.

CONTENTS

15.1 INTRODUCTION

High voltage direct current (HVDC) transmission is widely recognized as being advantageous for long-distance, bulk-power delivery, asynchronous interconnections and long submarine cable crossings. HVDC lines and cables are less expensive and have lower losses than those for 3-phase ac transmission.

Typical HVDC lines utilize a bipolar configuration with two independent poles and are comparable to a double circuit ac line. Because of their controllability HVDC links offer firm capacity without limitation due to network congestion or loop flow on parallel paths. Higher power transfers are possible over longer distances with fewer lines with HVDC transmission than with ac transmission. Higher power transfers are possible without distance limitation to HVDC cables systems using fewer cables than with ac cable systems due to their charging current.

HVDC systems became practical and commercially viable with the advent of high voltage mercury-arc valves in the 1950s. Solid-state thyristor valves were introduced in the late 1960s, leading to simpler converter designs with lower operation and maintenance expenses and improved availability. In the late 1990s a number of newer converter technologies were introduced permitting wider use of HVDC transmission in applications, which might not otherwise be considered. A list of HVDC projects currently in operation or under construction is given in Table 15-1.

TABLE 15-1 HVDC Project List

Name of HVDC system	Year commissioned/ upgraded/ retired	Nominal capacity (MW)	DC voltage (kV)	B-B line/ cable (km)	Location
Under Construction					
ESTLINK	2006	350	150	106	Estonia-Finland
BASSLINK	2005	500	400	360	Australia
NORNED	2007	600	500	580	Norway-Netherlands
THREE GORGES-SHANGHAI	2007	3000	±500	900	China
NEPTUNE	2007	600	500	102	U.S.A.
MISSION	2007	150	±21	B-B	U.S.A.
Operational					
VANCOUVER 1	1968	312	+260	74	Canada
VOLGOGRAD-DONBASS	1962	720	±400	470	Russia
SAKUMA	1965/1993	300	2 × 125	B-B	Japan
NEW ZEALAND HYBRID	1965/92	1240	+270/-350	612	New Zealand
PACIFIC INTERTIE	1970/84/89/02	3100	±500	1361	U.S.A.
NELSON RIVER 1	1973/93	1854	+463/-500	890	Canada
GOTLAND HVDC LIGHT	1999	50	±60	70	Sweden
DIRECTLINK	2000	3 × 60	±80	59	Australia
MURRAYLINK	2002	200	±150	176	Australia
CROSS SOUND	2002	330	±150	40	U.S.A.
TROLL	2004	2 × 40	±60	70	Norway
EEL RIVER	1972	320	2 × 80	B-B	Canada
VANCOUVER 2	1977	370	−280	74	Canada
DAVID A. HAMIL	1977	100	50	B-B	U.S.A.
SHIN-SHINANO 1	1977	300	125	B-B	Japan
SQUARE BUTTE	1977	500	±250	749	U.S.A.
CAHORA-BASSA	1978	1920	±533	1420	Mocambique-South Africa
C.U.	1979	1128	±411	702	U.S.A.
ACARAY	1981	50	26	B-B	Paraguay
INGA-SHABA	1982	560	±500	1700	Zaire
EDDY COUNTRY	1983	200	82	B-B	U.S.A.
CHATEAUGUAY	1984	2 × 500	2 × 140	B-B	Canada
BLACKWATER	1985	200	57	B-B	U.S.A.
HIGHGATE	1985	200	56	B-B	U.S.A.
MADAWASKA	1985	350	140	B-B	Canada
MILES CITY	1985	200	82	B-B	U.S.A.
OKLAUNION	1985	220	82	B-B	U.S.A.
BROKEN HILL	1986	40	2 × 17 (±8,33)	B-B	Australia
CROSS CHANNEL BP 1+2	1986	2000	±270	71	France-U.K.

TABLE 15-1 HVDC Project List (*continued*)

Name of HVDC system	Year commissioned/ upgraded/ retired	Nominal capacity (MW)	DC voltage (kV)	B-B line/ cable (km)	Location
IPP (INTERMOUNTAIN)	1986	1920	±500	784	U.S.A.
ITAIPU 1	1986	3150	±600	796	Brazil
ITAIPU 2	1987	3150	±600	796	Brazil
URUGUAIANAI	1987	54	18	B-B	Brazil-Uruguay
VIRGINIA SMITH	1987	200	50	B-B	U.S.A.
FENNO-SKAN	1989	572	400	234	Finland-Sweden
McNEILL	1989	150	42	B-B	Canada
SILERU-BARSOOR	1989	100	±200	196	India
VINDHYACHAL	1989	500	2 × 69.7	B-B	India
RIHAND-DELHI	1992	1500	±500	814	India
SHIN-SHINANO 2	1992	300	125	B-B	Japan
BALTIC CABLE	1994	600	450	255	Sweden-Germany
KONTEK	1995	600	400	171	Denmark-Germany
WELSH	1995	600	162	B-B	U.S.A.
CHANDRAPUR-RAMAGUNDUM	1997	1000	2 × 205	B-B	India
CHANDRAPUR-PADGHE	1998	1500	±500	736	India
HAENAM-CHEJU	1998	300	±180	101	South Korea
LEYTE-LUZON	1998	440	350	443	Philippines
VIZAG 1	1998	500	205	B-B	India
MINAMI-FUKUMITZU	1999	300	125	B-B	Japan
KII CHANNEL	2000	1400	±250	102	Japan
SWEPOL LINK	2000	600	450	230	Sweden-Poland
GRITA	2001	500	400	313	Greece-Italy
HIGASHI-SHIMIZU	2001	300	125	B-B	Japan
MOYLE INTERCONNECTOR	2001	2 × 250	2 × 250	64	Scotland-N.Ireland
TIAN-GUANG	2001	1800	±500	960	China
THAILAND-MALAYSIA	2001	600	±300	110	Thailand-Malaysia
EAST-SOUTH INTERCONNECTOR	2003	2000	±500	1400	India
RAPID CITY TIE	2003	2 × 100	±13	B-B	U.S.A.
THREE GORGES CHANGZHOU	2003	3000	±500	890	China
GUI-GUANG	2004	3000	±500	936	China
THREE GORGES-GUANGDONG	2004	3000	±500	900	China
LAMAR	2005	211	±63	B-B	U.S.A.
VIZAG 2	2005	500	±88	B-B	India
KONTI-SKAN 1 AND 2	1965/88/2005	740	±285	150	Denmark-Sweden
SACOI	1967/85/93	300	±200	385	Italy-Corsica-Sardinia
SKAGERRAK 1-3	1976/77/93	1050	250/350	240	Norway-Denmark
NELSON RIVER 2	1978/85	2000	±500	940	Canada
HOKKAIDO-HONSHU	1979/80/93	600	±250	167	Japan
VYBORG	1981/82/84/02	4 × 355	1 × 170 (±85)	B-B	Russia-Finland
GOTLAND II-III	1983/87	260	150	98	Sweden
QUEBEC-NEW ENGLAND	1986/90/92	2250	±500	1500	Canada-U.S.A.
GESHA	1989/90	1200	±500	1046	China
GARABI 1&2	2000/02	2000	±70	B-B	Argentina-Brazil
RIVERA		70		B-B	Uruguay
SASARAM	2002	500	205	B-B	India
Retired					
KINGSNORTH	1972/1987	640		82	England
DUERNROHR 1	1983/1997	550	145	B-B	Austria-Czech
ETZENRIHT	1993/1997	600	160	B-B	Germany-Czech
VIENNA SOUTH-EAST	1993/1997	600	145	B-B	Austria-Hungary

15.2 APPLICATIONS

The significant increase in HVDC transmission can be attributed to one or more of the following reasons:

Economical. HVDC transmission systems often provide a more economical alternative to ac transmission for long-distance, bulk-power delivery from remote resources such as hydroelectric developments, mine-mouth power plants, or generation from large-scale wind farms. Whenever long-distance transmission is discussed, the concept of "breakeven distance" frequently arises. This is where the savings in line costs and lower capitalized cost of losses offsets the higher converter station costs. A bipolar HVDC line uses only two insulated sets of conductors rather than three. This results in narrower right-of-way (ROW), smaller transmission towers, and lower line losses than with ac lines of comparable capacity. A rough approximation of the savings in line construction is 30%. Although breakeven distance is influenced by the costs of ROW and line construction with a typical value of 500 km, the concept itself is misleading because in many cases more ac lines are needed to deliver the same power over the same distance due to system stability limitations. Furthermore, the long-distance ac lines usually require intermediate switching stations and reactive power compensation. For example, the generator outlet transmission alternative for the ±250 kV, 500 MW Square Butte Project was two 345 kV series-compensated ac transmission lines. Similarly, the ±500 kV, 1600 MW Intermountain Power Project (IPP) ac alternative comprised two 500 kV ac lines. The IPP takes advantage of the double circuit nature of the bipolar line and includes a 100% short-term and 50% continuous monopolar overload. The first 6000 MW stage of the transmission for the Three Gorges Project in China would have required 5 × 500 kV ac lines as opposed to 2 × (±500) kV, 3000 MW bipolar HVDC lines (Fig. 15-1).

For underground or submarine cable systems there is considerable savings in installed cable costs and cost of losses with HVDC transmission. Depending on the power level to be transmitted, these savings can offset the higher converter station costs at distances of 40 km or more. Furthermore, there is a rapid drop-off in cable capacity with ac transmission over distance due to the reactive component of charging current. Although this can be compensated by intermediate shunt compensation for underground cables, it is not practical to do so for submarine cables. For a given cable conductor area, the line losses with HVDC cables, can be less than half those of ac cables. This is due to more conductors, reactive component of current, skin effect, and induced currents in the cable sheath and armor.

Functional. The controllability and asynchronous nature of HVDC transmission provides a number of advantages for certain transmission applications. HVDC transmission capacity is firm and utilization usually runs higher due to its controllability. This is because congestion or loop flow on parallel transmission paths does not result in schedules curtailments for transmission loading relief.

With a cable system, unequal loadings or risk of postcontingency overloads often results in use of a series-connected phase-shifting transformer. These potential problems do not exist with a controlled HVDC cable system.

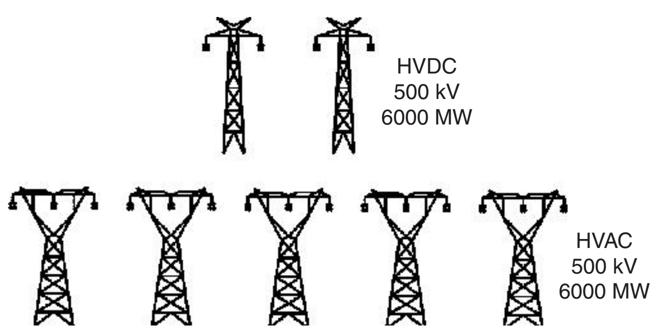


FIGURE 15-1 HVDC and EHV ac alternatives for first stage of three Gorges outlet transmission.

With HVDC transmission systems, interconnections can be made between asynchronous networks for more economic or reliable operation. The asynchronous interconnection allows interconnections of mutual benefit but provides a buffer between the two systems. Often these interconnections use back-to-back converters with no transmission line. The asynchronous links act as an effective "firewall" against propagation of cascading outages in one network from passing to another network. Many asynchronous interconnections exist in North America between the eastern and western interconnected systems, between the Electric Reliability Council of Texas (ERCOT) and its neighbors, that is, Mexico, Southwest Power Pool (SPP) and the western interconnect, and between Quebec and its neighbors, that is, New England and the Maritimes. The August 2003 northeast blackout provides an example of the firewall against cascading outages provided by asynchronous interconnections. As the outage propagated around the lower Great Lakes and through Ontario and New York, it stopped at the asynchronous interface with Quebec. Quebec was unaffected, the weak ac interconnections between New York and New England tripped, but the HVDC links from Quebec continued to deliver power to New England.

Environmental. HVDC allows delivery of more power over fewer lines with narrower ROW. This is especially important in trying to access diverse resources in remote locations where lines may pass through environmentally sensitive or scenic areas. There is no induction or alternating electromagnetic fields from HVDC transmission. There is no physical restriction limiting the distance for underground cables. Underground cables can be used on shared ROW with other utilities without impacting reliability concerns over use of common corridors. Lower cable losses improves efficiency and results in less heating in the earth.

15.3 HVDC FUNDAMENTALS

15.3.1 Converter Behavior and Equations

Conventional HVDC transmission schemes utilize line-commutated, current-source converters. Such converters require a synchronous voltage source in order to operate. The basic building block used for HVDC conversion is the 3-phase, full-wave bridge referred to as a 6-pulse or Graetz bridge (Fig. 15-2). The term 6-pulse is due to the characteristic harmonic ripple in the dc output voltage, which is at multiples of 6 times the fundamental frequency. Each 6-pulse bridge is comprised of 6 controlled switching elements or thyristor valves. Each valve comprises a number of series-connected thryristors to achieve the desired dc voltage rating.

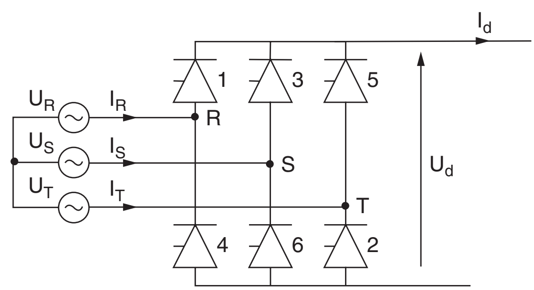


FIGURE 15-2 6-pulse bridge.

Converter dc output voltage is controlled by means of a delayed firing angle. Valve switching is synchronized to the ac source voltages via a phase-locked loop. The bridge is coupled to the ac bus via a converter transformer. Commutation of converter currents from one phase to another results in a converter voltage drop. Converter voltage drop is proportional to transformer reactance and current level I_d, resulting in a reduction of the dc voltage level U_d, due to commutation overlap u.

A set of equations has been derived to calculate U_d as a function of the phase voltages, the commutation reactance I_d, and the delay angle α. For rectifier operation converter polarity is positive, whereas for inverter operation it is negative bucking the direction of direct current flow. Equations describing inverter operation use extinction angle γ.

The direct voltage across the 6-pulse bridge is calculated by Eq. (15-1) for rectifier operation and Eq. (15-2) for inverter operation.

$$\frac{U_{dR}}{2} = U_{diOR}\cdot\left[\cos\alpha - (d_{xR} + d_{rR})\frac{I_d}{I_{dN}}\frac{U_{diORN}}{U_{diOR}}\right] - U_T \tag{15-1}$$

$$\frac{U_{dI}}{2} = U_{diOI}\cdot\left[\cos\gamma - (d_{xI} - d_{rI})\frac{I_d}{I_{dN}}\frac{U_{diOIN}}{U_{diOI}}\right] + U_T \tag{15-2}$$

The nominal relative inductive direct voltage drop is defined by Eq. (15-3), where X_t is the commutation reactance which includes the converter transformer reactance and any other reactances in the commutation circuit.

$$d_{xN} = \frac{3}{\pi}\frac{X_t\cdot I_{dN}}{U_{diON}} \tag{15-3}$$

The relative resistive direct voltage drop is defined by Eq. (15-4) where P_{cu} is the transformer and smoothing reactor load losses and R_{th} is current dependent voltage drop over the thyristors. The factor 2 is due to the fact that there are always two valves conducting at the same time.

$$d_r = \frac{P_{cu}}{U_{diON}\cdot I_{dN}} + \frac{2\cdot R_{th}\cdot I_{dN}}{U_{diON}} \tag{15-4}$$

The overlap angle for the rectifier and inverter are described by Eqs. (15-5) and (15-6), respectively.

$$\cos(\alpha + \mu_R) = \cos\alpha - 2\cdot d_{xNR}\frac{I_d}{I_{dN}}\frac{U_{diONR}}{U_{diOR}} \tag{15-5}$$

$$\cos(\gamma + \mu_I) = \cos\gamma - 2\cdot d_{xNI}\frac{I_d}{I_{dN}}\frac{U_{diONI}}{U_{diOI}} \tag{15-6}$$

The reactive power consumption for a 12-pulse converter (two 6-pulse converters with 30° shift in valve voltages) connected in series is calculated with Eq. (15-7).

$$Q_d = 2\cdot\chi\cdot I_d\cdot U_{diO} \tag{15-7}$$

where c is the overlap function described by Eq. (15-8) for rectified operation and Eq. (15-9) for inverter operation.

$$\chi = \frac{1}{4}\cdot\frac{2\cdot\mu + \sin 2\alpha - \sin 2(\alpha + \mu)}{\cos\alpha - \cos(\alpha + \mu)} \tag{15-8}$$

$$\chi = \frac{1}{4}\cdot\frac{2\cdot\mu + \sin 2\gamma - \sin 2(\gamma + \mu)}{\cos\gamma - \cos(\gamma + \mu)} \tag{15-9}$$

The relationship between the no-load phase-phase ac voltage on the valve side and the ideal no-load direct voltage is shown in Eq. (15-10). The rms value of the rated ac current on the valve side of the converter transformer is shown in Eq. (15-11). The total rated MVA of the 3-phase transformer group feeding the 6-pulse converter bridge is according to Eq. (15-12).

$$U_{vo} = \frac{U_{diO}}{\sqrt{2}}\cdot\frac{\pi}{3} \tag{15-10}$$

$$I_{vN} = \sqrt{\frac{2}{3}}\cdot I_{dN} \tag{15-11}$$

$$S_N = \sqrt{3}\cdot U_{vN}\cdot I_{vN} = \frac{\pi}{3}\cdot U_{diON}\cdot I_{dN} \tag{15-12}$$

Figure 15-3 illustrates the commutation process and its effect on valve currents and dc voltage due to delay angle and overlap. The solid upper envelope of the phase voltages is the voltage top of the bridge

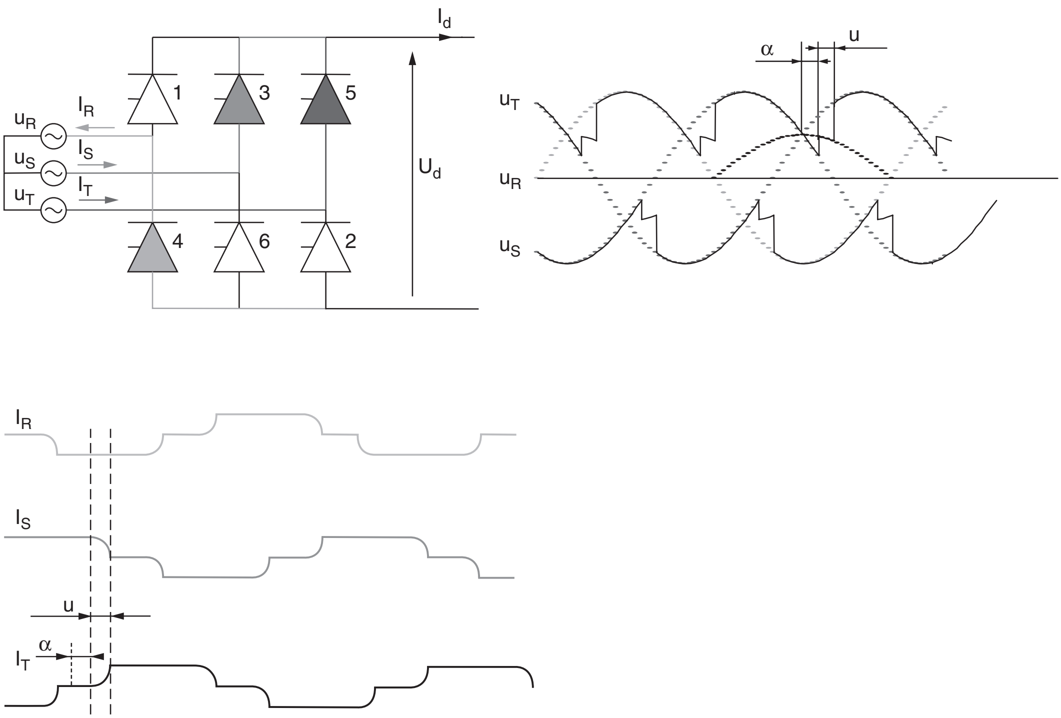


FIGURE 15-3 6-Pulse bridge commutation with delay angle and overlap.

with common valve cathodes, while the lower solid envelope is the voltage at the bottom of the bridge with the common valve anodes. The differential voltage across the bridge is the dc voltage U_d. The effect of the delay angle and commutation overlap on the dc voltage is evident. During commutation two valves in the same half bridge conduct simultaneously and the instantaneous voltage is half their sum.

The 6-pulse converter bridge can be used in rectifier operation with positive output voltage, $0 > \alpha < 90°$, converting ac to dc or in inverter operation with an output voltage that is negative with respect to the direction of dc current flow, $90 > \alpha < 180°$. By connecting two converters in series at opposite ends of a transmission line, one controlling dc voltage and the other controlling dc current, dc power transmission is achieved. The characteristic current harmonics ($f = 6n \pm 1$) are filtered on the ac side and the characteristic voltage harmonics ($f = 6n$) are filtered on the dc side to meet voltage distortion and telephone interference requirements.

The dc terminals of two 6-pulse bridges with ac voltage sources phase displaced by 30° can be connected in series for 12-pulse operation. In 12-pulse operation, the characteristic current and voltage harmonics have frequencies of $12n \pm 1$ and $12n$, respectively. The 30° phase displacement can easily be achieved by feeding one bridge through a transformer with a wye-connected secondary and the other transformer through a delta-connected secondary (Fig. 15-4). Most modern HVDC transmission

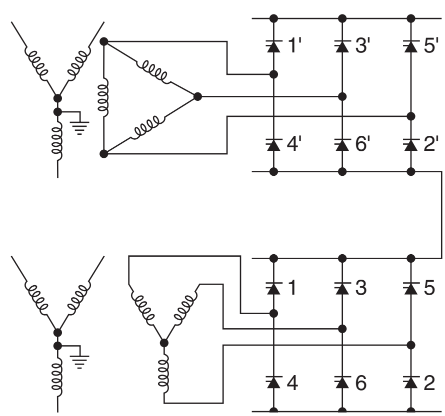


FIGURE 15-4 12-Pulse bridge.

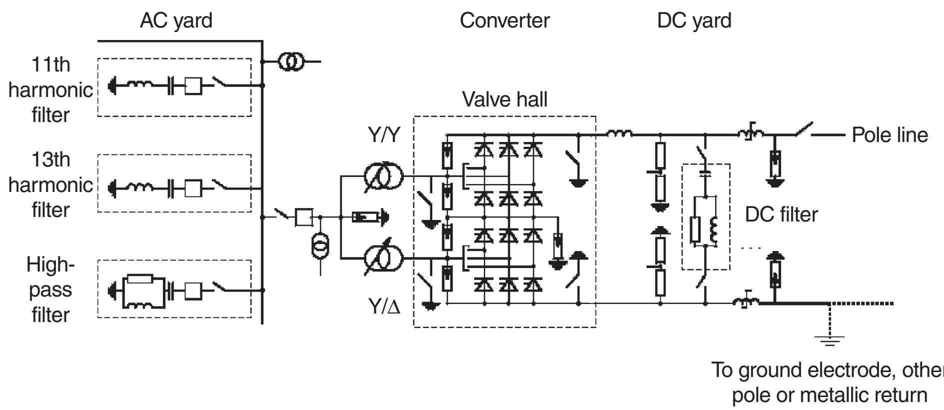


FIGURE 15-5 Simplified single line diagram for monopole.

schemes utilize 12-pulse converters to reduce the additional harmonic filtering requirements required for 6-pulse operation, for example, fifth and seventh on the ac side and sixth on the dc side. This is because although these harmonic currents still flow through the valves and the transformer windings, they are 180° out of phase and cancel out on the primary side.

15.3.2 Station Layout and System Configuration

A simplified single-line diagram for one pole with a 12-pulse converter is shown in Fig. 15-5. A CAD drawing and a photo of a monopolar converter station are shown in Figs. 15-6 and 15-7, respectively.

An HVDC converter station comprises the following major subsystems:

FIGURE 15-6 Monopolar converter station.

- Thyristor valves
- Converter transformers
- AC harmonic filters
- DC harmonic filters
- Valve cooling
- Control and protection
- Auxiliary power
- Valve hall building

The converter station layout depends on a number of factors such as the station configuration, that is, monopolar (Fig. 15-8), bipolar (Fig. 15-9) or back-to-back asynchronous tie (Fig. 15-10), valve design, ac system interconnection, filtering requirements, reactive power compensation requirements, land availability, and the local environment. In most cases, the thyristor valves are air-insulated, water-cooled, and enclosed in a converter building often referred to as a valve hall. For back-to-back ties with their characteristically low dc voltage, thyristor valves can be housed in prefabricated electrical enclosures in which case a valve hall is not required.

To obtain a more compact station design and reduce the number of insulated high voltage wall bushings, converter transformers are often placed adjacent to the valve hall with valve winding bushings protruding through the building walls for connection to the valves. Double or quadruple valve structures housing valve modules are used within the valve hall. Valve arresters are located immediately adjacent to the valves. Indoor motor-operated grounding switches are used for personnel safety

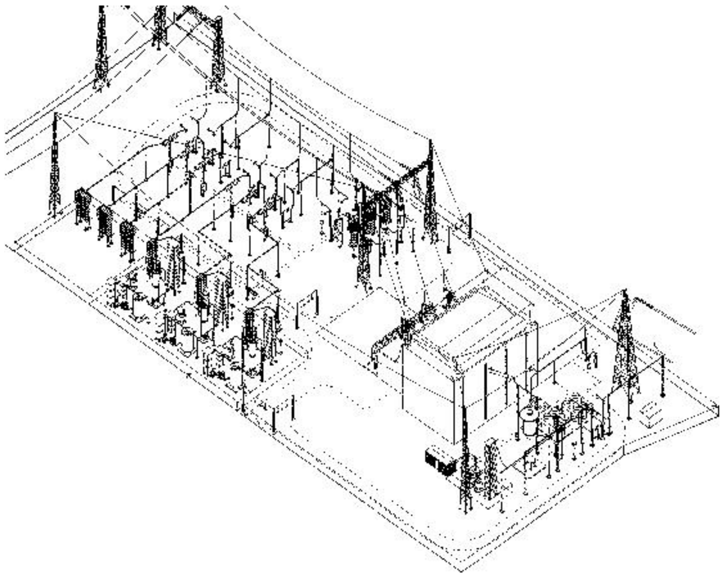


FIGURE 15-7 CAD drawing of monopolar converter station.

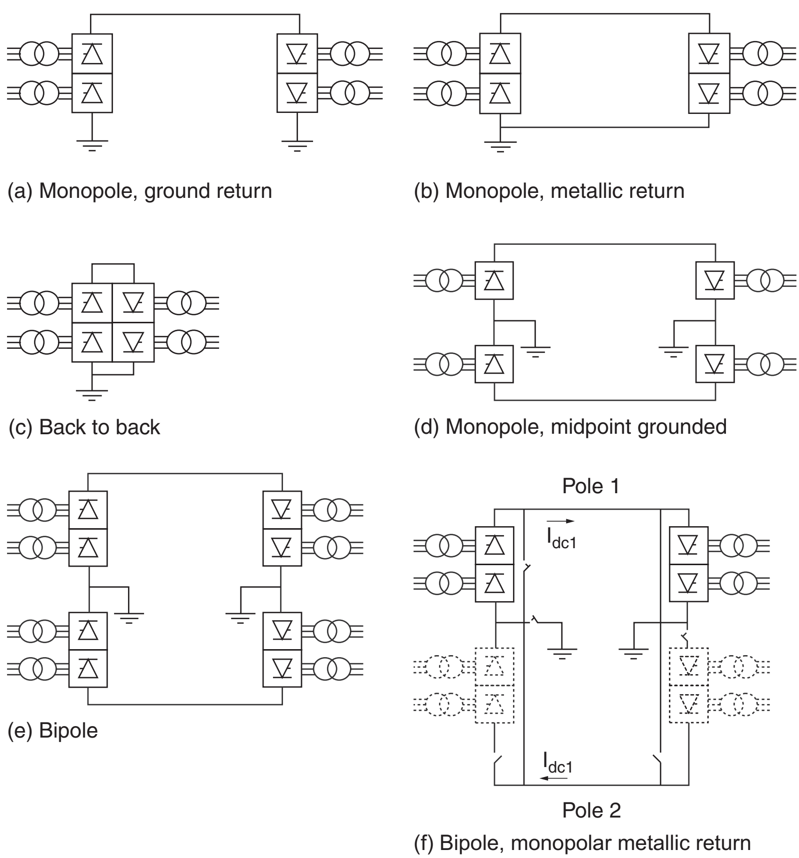


FIGURE 15-8 HVDC operating configurations/modes.

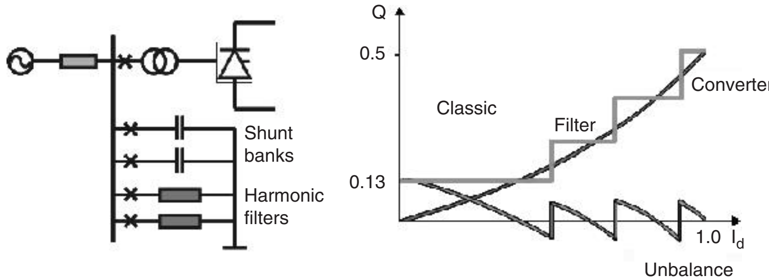


FIGURE 15-9 Reactive power balance.

during maintenance. Closed loop valve cooling systems are used to circulate the cooling medium through the indoor thyristor valves with heat transfer to dry coolers or evaporative cooling towers located outdoors.

Monopolar systems with ground return are the simplest and least expensive systems for moderate power transfers since only two converters and one insulated cable or line conductor is required. Such systems are commonly used with low voltage electrode lines and sea electrodes to carry the return current in submarine cable crossings.

In some areas conditions are not conducive to monopolar earth or sea return. This could be the case areas in heavily congested areas, fresh water cable crossings, or areas with high earth resistivities. In such cases a metallic neutral or low voltage cable is used for the return path and the dc circuit uses a simple ground local ground reference.

Back-to-back stations are used for interconnection of asynchronous networks and use ac lines to connect on either side. In such systems power transfer is limited by the relative capacities of the adjacent ac systems at the point of coupling.

As an economic alternative to a monopolar system with metallic return, the midpoint of a 12-pulse converter can be connected to earth directly or through an impedance and two half voltage cables or line conductors can be used. The converter is only operated in 12-pulse mode, so there is no earth current.

The most common configuration for modern overhead HVDC transmission lines is bipolar with a single 12-pulse converter for each pole at each terminal. This gives two independent dc circuits each capable of half capacity. For normal balanced operation there is no earth current. Monopolar earth return operation, often with overload capacity, can be used during outages of the opposite pole.

Earth return operation can be minimized during monopolar outages by using the opposite pole line for metallic return via pole/converter bypass switches at each end. This requires a metallic-return transfer breaker in the ground electrode line at one of the dc terminals to commutate the current from the relatively low resistance of the earth into that of the dc line conductor. Metallic return operation

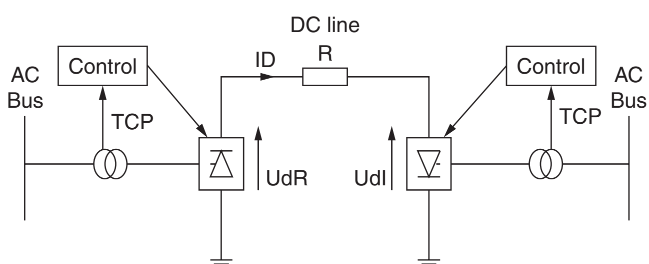


FIGURE 15-10 HVDC control system.

capability is provided for most dc transmission systems. This is not only effective during converter outages but also during line insulation failures where the remaining insulation strength is adequate to withstand the low resistive voltage drop in the metallic return path.

15.3.3 Reactive Power Compensation

As shown by Eqs. (15-7) through (15-9) in Sec. 15.3.1, HVDC conversion with line-commutated converters demands reactive power from the ac network at each HVDC terminal. The reactive power demand is a function of the firing angle in rectifier operation and extinction angle in inverter operation, the direct current and the overlap angle. The overlap angle is a function of the ac commutating voltage, the commutation reactance, and the dc current. As a rough approximation nominal reactive power demand at each terminal is about half the active power transfer.

The total reactive power produced by all the ac harmonic filters at each terminal is usually in the range of 30% to 40% of the converter rating. The filters therefore provide most of the reactive power compensation to meet the converter reactive power demand. The remaining reactive power necessary at the higher power levels can be provided from shunt capacitor banks, synchronous condensers, and static var compensators or nearby generation. Any reactive power mismatch must be provided or absorbed by the local ac system. Figure 15-9 shows the reactive power demand of a converter station, the reactive power from the filters, and the reactive power exchange with the ac network as a function of power transfer.

With weaker ac networks, that is, networks where the 3-phase symmetrical short circuit capacity is low compared to the rating of the dc converter station, various system constraints impact the reactive power compensation. With weaker systems, the size of the reactive power compensation elements may need to be reduced due to the voltage change on switching and the allowable reactive power exchange with the ac network. This may mean that filter banks may have to be subdivided with smaller branches. Sometimes, the minimum filtering requirements, for example, those at low power, exceed the reactive power demand of the converters, and shunt reactors are also required to absorb the excess vars from the filters.

15.3.4 Control and Operation of HVDC Links

The fundamental objectives of an HVDC control system are:

- To control basic system quantities such as dc line current, dc voltage, and transmitted power accurately and with sufficient speed of response
- To maintain adequate commutation margin in inverter operation so that the valves can recover their forward blocking capability after conduction before their voltage polarity reverses
- To control higher level quantities such as frequency in isolated mode or provide power oscillation damping to help stabilize the ac network
- To compensate of loss of a pole, a generator, or ac transmission circuit by rapid readjustment of power
- To ensure stable operation with reliable commutation in the presence of system disturbances
- To minimize system losses and reactive power consumption
- Ensure proper operation with fast and stable recoveries during system faults and disturbances

With HVDC transmission one terminal sets the dc voltage level, while the other regulates the dc current by controlling its output voltage relative to that maintained by the voltage-setting terminal. Since the dc line resistance is low, large changes in current and hence power can be made with relatively small changes in firing angle. Two independent methods exist for controlling the converter dc output voltage. These are (1) by changing the ratio between the direct voltage and the ac voltage by varying the delay angle α or (2) by changing the converter ac voltage via load tap changers (LTC) on the converter transformer. Although the former method is rapid, the latter method is slow due to

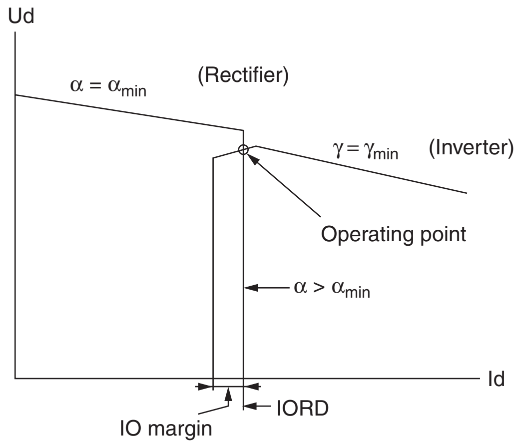


FIGURE 15-11 Static operating characteristics.

the limited speed of response of the LTC. Use of high delay angles to achieve a larger dynamic range, however, increases the converter reactive power consumption. To minimize the reactive power demand while still providing adequate dynamic control range and commutation margin, the LTC is used at the rectifier terminal to keep the delay angle within its desired steady-state range, for example, 13° to 18°, and at the inverter to keep the extinction angle γ within its desired range, for example, 17° to 20°, if the angle is used for dc voltage control or maintain rated dc voltage if operating in minimum commutation margin control mode.

Cooperation between the two terminals allows for efficient operation and provides for backup control modes for abrupt changes to the system voltages during disturbances. The converter control system at each terminal provides a static control characteristic. The intersection of the static control characteristics at the rectifier and inverter terminals determines the operating point. With the rectifier operating in constant current control and the inverter in constant angle control, as shown in Fig. 15-11, presents a stable operating point.

Each converter terminal is equipped with a closed loop current control or current control amplifier (CCA) as shown in Fig. 15-12. The backup current regulator at the inverter comes into effect when the rectifier ac voltage is suddenly reduced, forcing the rectifier characteristic down resulting in a new operating point with the rectifier minimum firing angle setting the dc voltage and the inverter current order setting the current. This shift in operating point is referred to a mode shift. A dc voltage regulator may also be used with or without current compounding to achieve a positive slope at the inverter with minimum extinction angle or commutation margin as a backup. A mode shift can also occur for a sudden increase in inverter ac voltage if operating in constant extinction angle control.

Other control functions are needed to synchronize the valve firing to the ac system commutation voltages, to clear and recover from dc line faults, to translate the alpha orders to firing pulses and

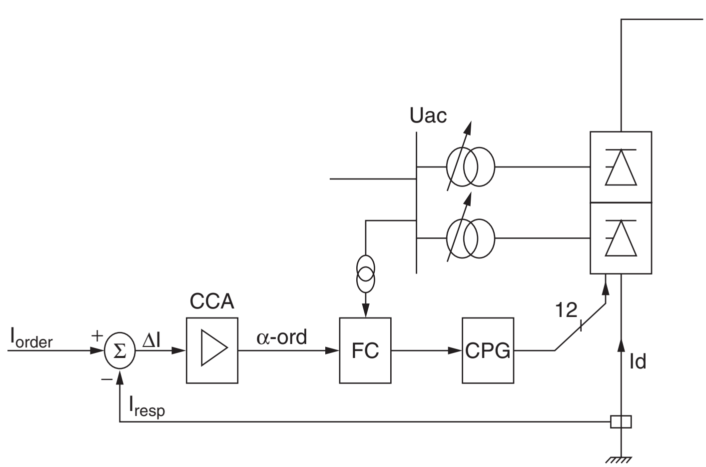


FIGURE 15-12 Closed loop current control system.

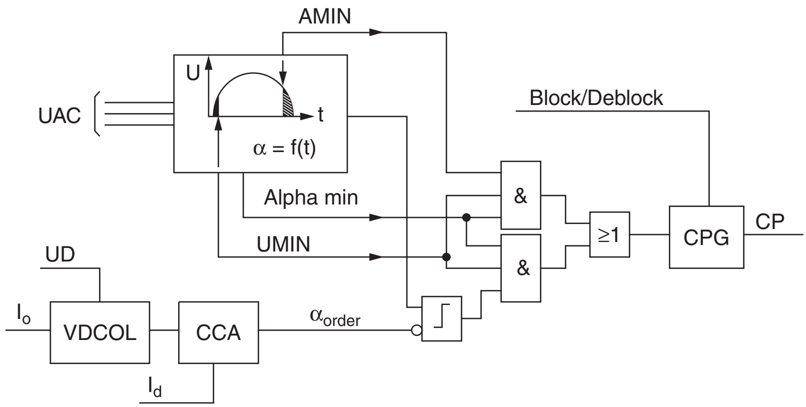


FIGURE 15-13 Converter firing control.

distribute them to the high-voltage valves, to minimize the reactive power consumption and achieve stable recoveries from large signal disturbances and faults in the ac network. Figure 15-13 shows these basic functions in the converter firing control (CFC).

The current order I_o is received from the pole power control. If the dc voltage is very low as during faults, the current order is limited by the voltage-dependent current order limiter, VDCOL. The alpha firing order is then limited as to its minimum and maximum value and minimum valve firing voltage (UMIN) in the converter firing control. Alpha min is used in inverter operation to prevent firing in rectifier operation. Minimum commutation margin control is used in inverter operation to maintain the minimum voltage time area to ensure successful recovery of forward blocking capability after valve conduction.

Figure 15-14 shows the static characteristics of the rectifier and inverter with addition of the VDCOL. The VDCOL acts to limit the dc current order below its normal set point if the dc current is above its break point and the dc voltage is lower than its break point. Taking into account dynamic performance, the current limitation is very fast acting during decreasing voltage due to faults, while the recovery is slower upon system voltage recovery depending on ac system strength or ability to deliver reactive power to the converter during recovery.

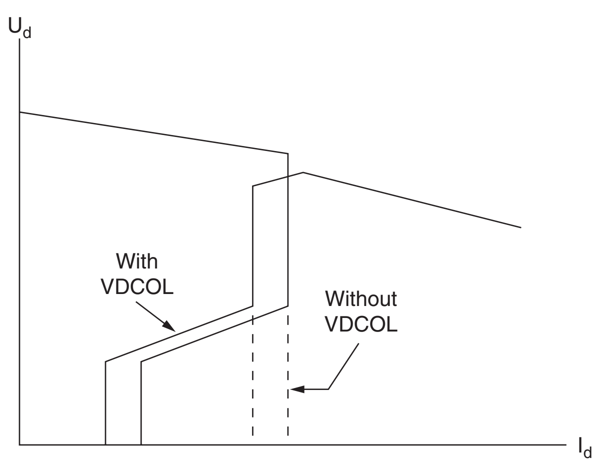


FIGURE 15-14 U_d-I_d characteristics with VDCOL.

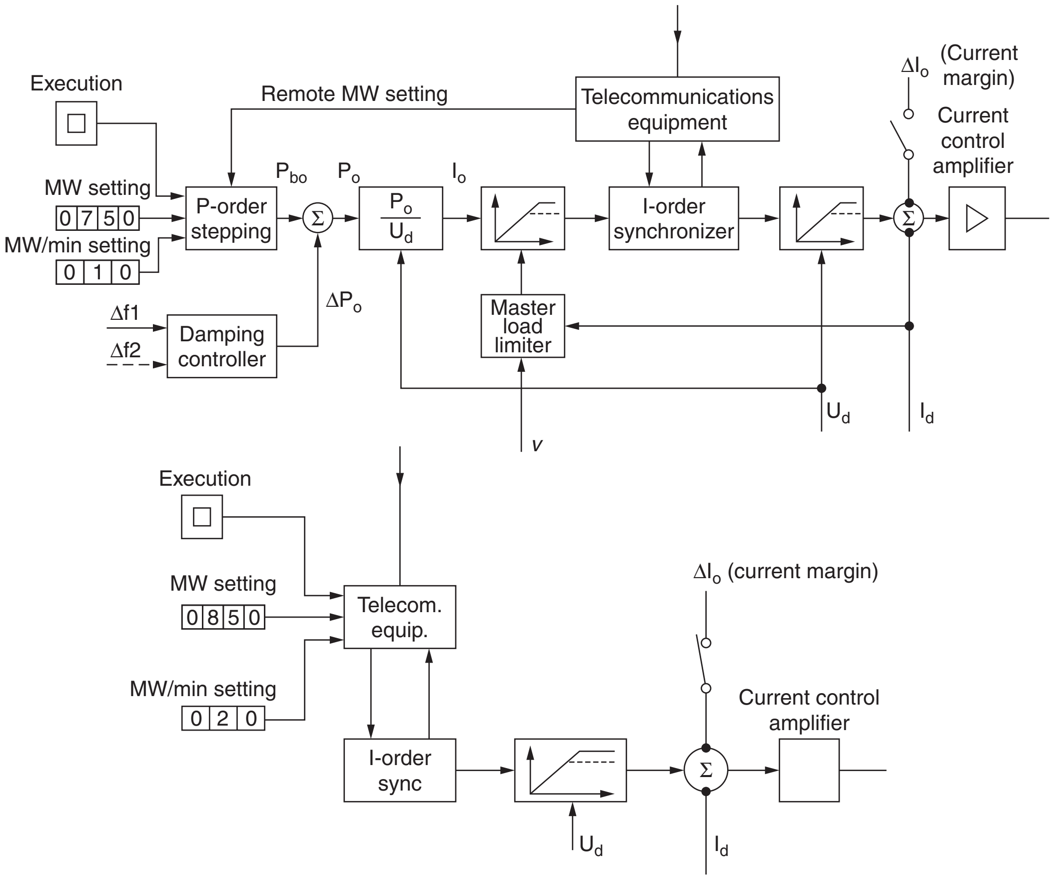


FIGURE 15-15 Master power control and current order synchronization.

The fundamental control functions described in the previous paragraphs are applied at the pole level and are independent of those on the other pole in a bipolar system. Coordination of the current orders between the terminals is required during ramping of the dc power during schedule changes. This is done during normal operation with secure communications between the terminals. Backup control strategies have been developed for communications outages. In a bipolar system, a master control is used for coordinated schedule changes and calculation of the current orders for each pole. The master control is used for compensation for loss of a pole by doubling the current order on the remaining pole subject to the equipment ratings. Figure 15-15 shows the current order coordination between the two terminals. For bipolar operation, the voltage fed to the power controller is the bipolar voltage assuring equal current orders to each pole. Upon loss of a pole this voltage is cut in half. Normally, the master control is intentionally slow being only used for schedule changes. For loss of a pole, however, its response time is fast. The master control can also handle supplemental control functions such as power oscillation damping and frequency control. Synchronization of the current order is such that the current margin is maintained.

15.3.5 Multiterminal Operation

The same control principles used for two-terminal operation can be applied to multiterminal operation with one terminal being assigned to voltage control, while the other terminals control their

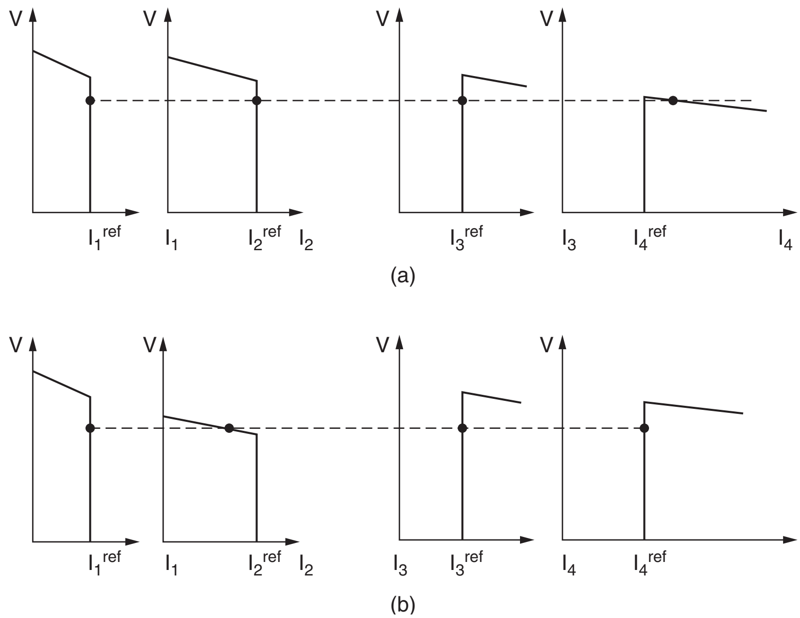


FIGURE 15-16 Static characteristics for 4-terminal HVDC system illustrating mode shift from inverter 4 (upper set) to Rectifier 2 (lower set) due to depressed ac voltage at Rectifier 2.

respective dc current orders (Fig. 15-16). The master control must also ensure that the sum of the rectifier current orders equals the sum of the inverter current orders on a per pole basis during all operating conditions. If one of the terminals is limited or tripped, the residual mismatch is allocated among the remaining stations according to prioritized distribution factors to ensure that Kirchoff's law is met. If the tripped station is the voltage setting terminal (VST), one of the remaining stations must be assigned to voltage control. The same method for clearing dc line faults, force retard of the rectifier(s) to invert off the dc current, can be used along with fast-acting pole-isolating switches which in turn can be used to isolate a faulty terminal without using special purpose dc breakers.

15.3.6 Economics and Efficiency

The following factors influence the optimum solution for HVDC transmission systems:

- Power transfer requirements
- Transmission distance
- Capitalized cost of losses
- System configuration, that is, bipolar, monopolar, back-to-back OVHD line or cable system
- System connection voltages
- Relative system strength
- Reactive compensation requirements
- Environmental conditions
- Future expandability
- Transformer transport limitations

There is an economy of scale for HVDC transmission. It would cost less per kilowatt to transfer 3000 MW a distance of 800 km at ± 500 kV than it would to transfer 1000 MW. It would cost less

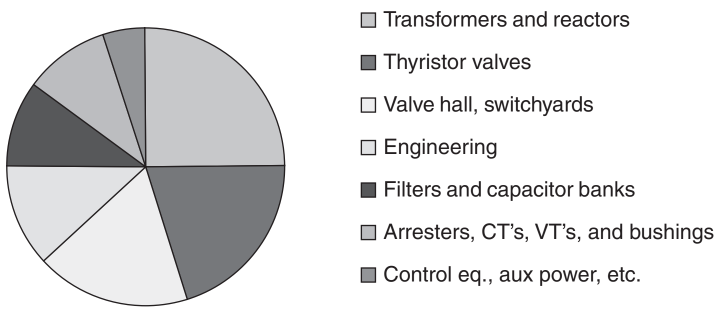


FIGURE 15-17 Terminal cost.

per kilowatt to transfer 600 MW over a monopolar submarine cable system than it would to transfer the same power on a 2-pole cable system with each pole rated at half the capacity. A 550-MW back-to-back asynchronous link would cost less per kilowatt than a 150-MW link.

HVDC applications at locations with relatively low short circuit capacities typically cost more per kilowatt due to constraints on reactive power compensation and dynamic overvoltage mitigation measures. A typical terminal cost breakdown of an HVDC transmission system for an OVHD line is shown in Fig. 15-17.

15.4 ALTERNATIVE CONFIGURATIONS

15.4.1 Capacitor-Commutated Converters

Converters with series capacitors connected between the valves and the transformers were introduced in the late 1990s for weak-system back-to-back applications. These converters are referred to as capacitor-commutated converters (CCC). The series capacitor provides some of the converter reactive power compensation requirements automatically with load current and provides part of the commutation voltage improving voltage stability. The overvoltage protection of the series capacitors is simple since the fault currents are limited by the impedance of the converter transformers. The CCC configuration allows higher power ratings in areas where the ac network is close to its voltage stability limit. The asynchronous Garabi interconnection between Brazil and Argentina consists of 4 × 550 MW parallel CCC links. The Rapid City Tie between the eastern and western interconnected systems consists of 2 × 100 MW parallel CCC links (Fig. 15-18). Both installations use a modular design with converter valves located within prefabricated electrical enclosures.

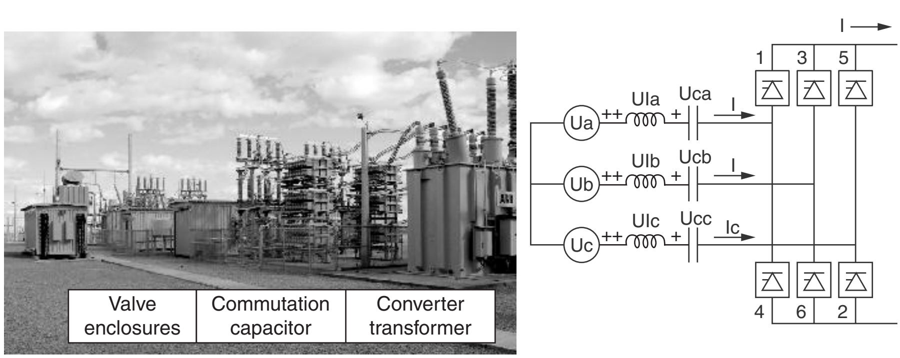


FIGURE 15-18 Rapid City Tie with modular 2 × 100 MW capacitor commutated converters.

15.4.2 Grid Power Flow Controller

A variation of the line-commutated design using a single 6-pulse converter has been used for a small back-to-back tie application. The term grid power flow controller (GPFC) has been used to describe this system design. By using a 6-pulse converter, there is no need for a second transformer secondary connection to obtain the requisite 30° phase displacement for 12-pulse operation. More ac harmonic filtering in the form of fifth and seventh branches is required, however. By using a 6-pulse converter and connecting the filters on the valve side, a simpler transformer connection can be utilized for matching the system voltage and blocking zero-sequence currents from flowing into the ac network. The ungrounded system has a large zero-sequence third order harmonic voltage component, however, appearing on the ungrounded neutrals and on the dc pole voltages, which increases the insulation levels. Despite using only one 6-pulse converter, the same number of series-connected thyristors is needed for the same dc voltage level.

15.4.3 Variable Frequency Transformer (VFT)

A technology that competes with HVDC for small capacity back-to-back ties in the 100 MW range was introduced in the early 2000s. A variable frequency transformer (VFT) is a machine rotating at the slip frequency between the two networks with high current between the rotor and stator passing through slip rings. The angle of the rotor is positioned to achieve a scheduled power flow by means of dc drives. The machine is connected to the network via step-up transformers. The reactive power demands of the VFT must be supplied by mechanically switched capacitor banks. Power control is slow due to having to move the inertia of the rotor, so it cannot respond quickly to a trip of generation on one the isolated network, for example. It cannot respond rapidly to variations in frequency or phase angle in the network so there will be inadvertent flow for fast variations. The VFT and its transformers provide an impedance, albeit a high one of around 40%, between the two networks. Therefore, the VFT will act as a voltage divider for faults in the network. This means that reactive power will be drained from one network due to a fault in the other. Losses of the VFT are higher than those for conventional HVDC.

15.5 STATION DESIGN AND EQUIPMENT

15.5.1 Thyristor Valves

For HVDC conversion, the thyristor valve must perform the following functions:

- Sequentially connect selected ac phases to the dc system per control pulses
- Conduct high current with low forward drop
- Block high voltages in both the forward and reverse directions
- Controllable and self monitoring
- Even voltage distribution and current turn-on
- Damp switching transients
- Fault tolerant and robust
- Accommodate cooling medium in high voltage environment

Thyristor valves are built up of series-connected thyristor modules and saturable reactors to limit valve turn-on di/dt. Each module contains a number of series-connected thyristors mounted on heat sinks. Each thyristor level is paralleled by an RC network for even voltage distribution and damping of commutation overshoots. Voltage measurement across each thyristor level is provided for thyristor monitoring, forward protection, and recovery protection.

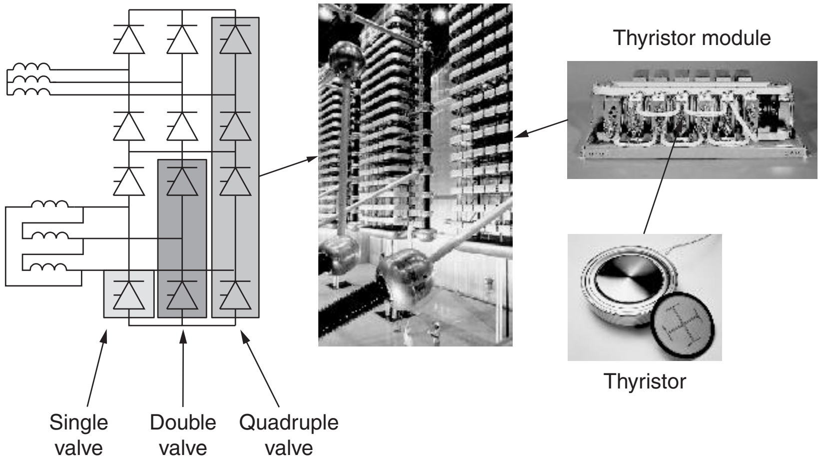


FIGURE 15-19 12-Pulse quadruple thyristor valve arrangement.

Each thyristor is coupled to the valve firing control at ground potential by means of two fiber optic links, one to carry valve trigger pulses to the thryistor gate circuit and the other for thyristor monitoring. Two types of thyristor triggering are used, electrically triggered thyristors (ETT) and light-triggered thyristors (LTT). Both triggering methods require voltage measurement at each thyristor level for monitoring and protection. ETT derives energy for gating from the RC damping circuit and gating is initiated by trigger pulses generated by light-emitting diodes. LTT thyristors have an optical turning-on region integrated on the thyristor wafer itself and use higher-power trigger pulses provided by laser diodes. Each thyristor level is equipped with forward protection which gates the thyristor on if the forward blocking voltage becomes too high due to, for example, absence of a trigger pulse. In inverter operation, during the thyristor recovery time after conduction, the forward protection level can be temporarily lowered. This is called recovery protection. ETT permits recovery protection to be implemented independently at the individual thyristor level (Fig. 15-19).

15.5.2 Converter Transformers

Converter transformers are the link between the ac and dc systems. They provide isolation between the two systems, preventing dc voltage and current from reaching the ac system. They also provide the phase displacement necessary for 12-pulse operation through wye- and delta-valve winding connections. Converter transformers have regulating windings with load-tap changers to maintain the ac voltage and converter firing angle within a narrow band across the entire converter operating range. Converter transformer impedance also limits the valve short-circuit levels to within their handling capability. As shown by Eq. 15-12, the 3-phase rating of the converter transformer for a 6-pulse bridge is proportional to U_{diON} and I_{dN}.

Converter transformer losses are those due to the fundamental frequency of load current plus those due to harmonics. The insulation design for converter transformers must take into account the direct voltage stresses superimposed on the normal ac voltage stresses. The ac stresses distribute as it would in a capacitive network while the dc voltage stresses distribute as according to a resistive network.

Transformer design depends on the bridge rating and type of converter connection and takes into account spare parts requirements and transport restrictions. For a small back-to-back, for example, a

3-phase bank with double secondary (wye and delta) may be used, that is, nine windings on a single core structure in a common tank for each 12-pulse converter bridge. For larger converters, three, single-phase transformers with double secondary windings may be used for each 12-pulse bridge. For the largest converter ratings where there may be some transport limitations, single-phase, two-winding transformers may be used, that is, six transformers per 12-pulse bridge (Fig. 15-20).

FIGURE 15-20 Single-phase, three-winding converter transformer for a 3100 MW bipole.

15.5.3 Smoothing Reactor

A smoothing reactor is connected in series with the converter on the dc side to reduce the harmonic ripple in the dc current as well as reduce transient currents during faults. The smoothing reactor also protects the converter valves from voltage surges coming in on the dc line. The dc smoothing reactor together with shunt-connected dc filters serve to limit telephone interference disturbing currents from flowing on the dc line. Most smoothing reactors are air-core, naturally air-cooled.

15.5.4 AC Filters

Converters inject harmonic currents into the ac network. AC filters are used to prevent these harmonic currents from flowing into the ac network impedance causing voltage distortion and induced telephone interference in the audible frequency range. AC filters provide a low-impedance path to ground at the harmonic frequencies. The ac filter comprises high-voltage capacitor banks and lower-voltage reactors, resistors, and capacitors, which together form a circuit tuned to the characteristic harmonic(s). The lower-order filters are single- or double-tuned, band-pass filters, while the higher harmonics are often taken care of by high-pass filters (Fig. 15-21).

AC harmonic filter design involves calculating the harmonic currents generated and estimating harmonic impedance characteristics of the ac network across the whole range of operating conditions and tolerances. A filter design is then developed to meet the required performance requirements. Filter components are then rated with an adequate margin the particular application.

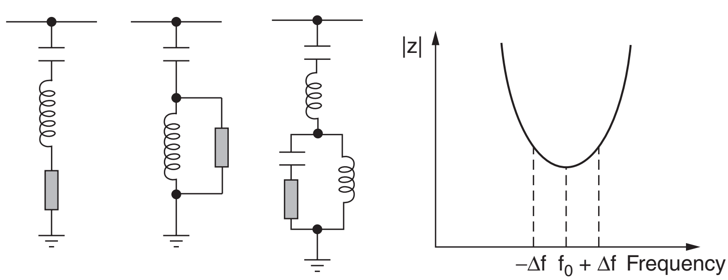


FIGURE 15-21 (*a*) Bandpass filter, (*b*) highpass filter, (*c*) double bandpass filter, (*d*) impedance vs. frequency.

The most common filter performance criteria are individual and total harmonic voltage distortion, D_T and D_h, and weighted telephone interference factor (TIF), calculated as follows:

$$D_h = 100 \times V_h/V_1$$

$$D_T = \left(\sum_{h=2}^{49} D_h^{\,2}\right)^{1/2}$$

$$\text{TIF} = \left[\sum_{h=2}^{49}\left(F_h \cdot \frac{V_h}{V_1}\right)^2\right]^{1/2}$$

15.5.5 DC Filters

Filters are required on the dc side for dc to limit interference with communication circuits, which are inductively coupled to the dc line, for example, parallel telephone lines. The design criterion for dc harmonic filters is a function of relating to the flow of harmonic currents at any point along the dc line to the interference with adjacent telephone lines. Significant parameters are the relative location of telephone lines with respect to the dc line, their shielding, the presence of any ground wires, and the earth's resistivity. This criterion is typically expressed as equivalent disturbing current I_{eq}. Disturbance levels are lower in normal balanced bipolar mode, due to cancellation effects, than in monopolar mode.

DC filter design must take into account the entire dc network with all harmonic sources and operating modes. DC harmonic filters consist of band-pass and high-pass filters connected in shut outside the smoothing reactor. Many modern HVDC links use a single 12th harmonic band-pass filter on each pole with active filtering for the higher order harmonics (Fig. 15-22). Active filtering consists of measuring the actual dc-side harmonics from the converter and counter-injecting the same amount with opposite polarity.

15.5.6 Power Line Carrier (PLC) Filters

Commutation in HVDC converters discharges stray capacitances and generates electrical noise at the lower end of the power line carrier spectrum (PLC), that is, strongest at 30 to 70 kHz. This noise may pass onto the interconnecting ac and dc lines. Where low-level carriers exist at the lower end of the PLC spectrum, filters may be required.

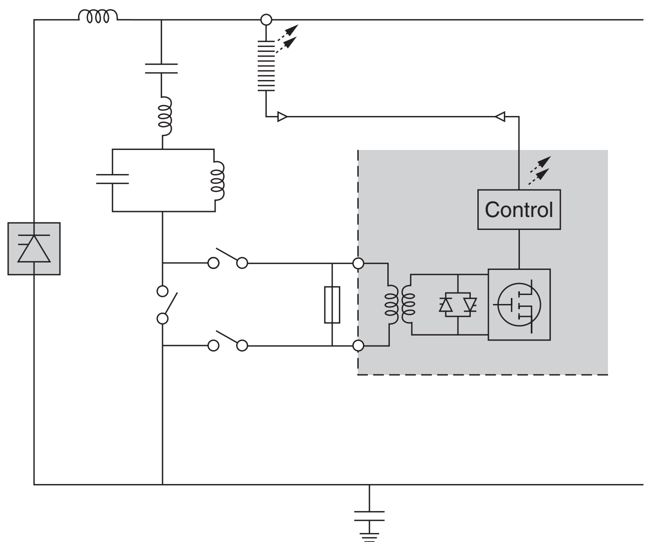


FIGURE 15-22 Active dc harmonic filter.

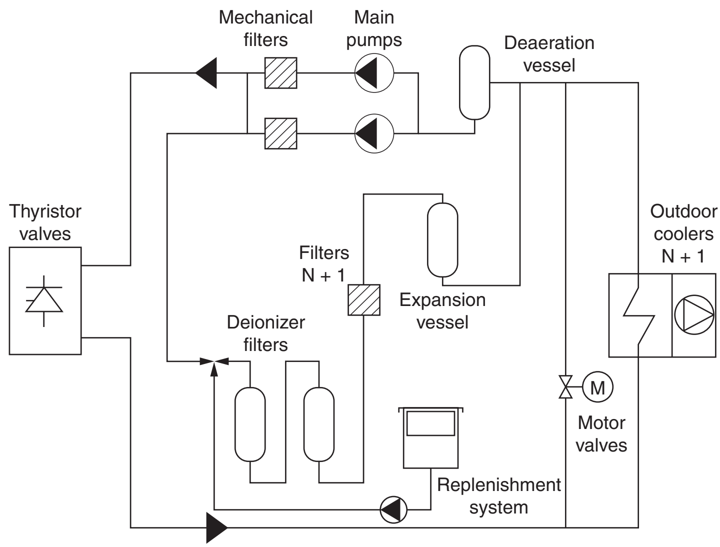


FIGURE 15-23 Closed-loop water cooling system.

15.5.7 Valve Cooling System

Thyristor valves must be cooled to avoid too high thyristor junction temperatures and to dissipate heat from the valve damping circuits and reactors. Valve cooling is accomplished by a deionized water loop circulating via insulated tuning to the individual thyristor heat sinks. Waste heat is passed to outdoor liquid-to-air coolers. Redundant variable speed pumps and coolers fed from redundant power supplies are used for reliability, availability, and ease of maintenance (Fig. 15-23).

15.5.8 Reliability and Availability

To meet high levels of reliability and availability plus facilitate ease of maintenance, redundancy is commonly used in HVDC converter station design. Typical guaranteed unavailability values are 0.5% for forced outages and 1.0% for scheduled outages.

Redundant series-connected thyristor levels are used in the valves. The failure mode is short circuit of the thyristor, so operation can continue until a convenient time for restoring full redundancy. Redundant cooling pumps and cooler units are used. Use of redundant control and protection systems is often used. For major main circuit components, spare parts are provided at site to minimize the time for replacement.

15.6 VOLTAGE SOURCE CONVERTER (VSC) BASED HVDC TRANSMISSION

15.6.1 System Characteristics

Conventional HVDC transmission employs line-commutated, current-source converters with thyristor valves. These converters require a relatively strong synchronous voltage source in order to commutate. The conversion process demands reactive power from filters, shunt banks, or series capacitors, which are an integral part of the converter station. Any surplus or deficit in reactive power must be

TABLE 15-2 HVDC VSC Projects Listing

Project	Year commissioned	Power rating, MW	DC voltage, kV	Cable, km	Location
Hellsjon	1997	3	± 10	10	Sweden
Gotland Light	1999	50	± 80	70	Sweden
Direct Link	2000	3 × 60	± 80	65	Australia
Tjaerborg	2000	7.2	± 9	4.4	Denmark
Cross Sound Cable	2002	330	± 150	40	United States
Murraylink	2002	200	± 150	180	Australia
Troll Offshore	2005	2 × 42	± 60	70	Norway
Estlink	2006	350	± 150	105	Estonia/Finland

accommodated by the ac system. This difference in reactive power needs to be kept within a given band to keep the ac voltage within the desired tolerance. The weaker the system or the further away from generation, the tighter the reactive power exchange must be to stay within the desired voltage tolerance.

HVDC transmission using voltage-source converters (VSC) with pulse-width modulation (PWM) was introduced as HVDC Light® in the late 1990s by ABB. These VSC-based systems are force-commutated with insulated-gate bipolar transistor (IGBT) valves and solid-dielectric, extruded HVDC cables (Table 15-2).

HVDC transmission and reactive power compensation with VSC technology has certain attributes which can be beneficial to overall system performance. VSC converter technology can rapidly control both active and reactive power independently of one another. Reactive power can also be controlled at each terminal independent of the dc transmission voltage level. This control capability gives total flexibility to place converters anywhere in the ac network since there is no restriction on minimum network short-circuit capacity. Forced commutation with VSC even permits black start, that is, the converter can be used to synthesize a balanced set of 3-phase voltages like a virtual synchronous generator. The dynamic support of the ac voltage at each converter terminal improves the voltage stability and increases the transfer capability of the sending and receiving end ac systems.

15.6.2 Applications

The aforementioned attributes of VSC-based HVDC transmission makes it especially suitable in certain applications. These applications are summarized as follows:

Underground Cable. HVDC cable systems do not face the distance limitations or suffer the higher losses of ac cable systems. Therefore, long-distance HVDC cable transmission is possible. Extruded HVDC cables are lighter, more flexible, and easier to splice than the mass-impregnated, oil-paper cables (MIND) used for conventional HVDC transmission, thus making them more conducive for land cable applications where transport limitations can drive up costs. The lower cost cable installations made possible by the extruded HVDC cables makes long-distance underground transmission economically feasible for use in areas with ROW constraints.

Power Supply to Insular Load. Forced-commutation, dynamic voltage control, and black-start capability allow VSC HVDC transmission to serve isolated loads on islands over long-distance submarine cables without any need for running expensive local generation.

Offshore. The VSC transmission is compact and can feed production or transportation loads on offshore oil or gas platforms from shore. This can eliminate the need for more expensive, less efficient, or higher emission offshore power production. The VSC converters can operate at variable frequency to more efficiently drive large compressor or pumping loads using high-voltage motors.

Asynchronous Interconnections. Interconnections between asynchronous networks are often at their periphery where the networks tend to be weak relative to the desired power transfer. The dynamic voltage support and improved voltage stability offered by VSC-based converters permits higher power transfers without as much need for ac system reinforcement. The VSC converters do not suffer commutation failures allowing fast recoveries from nearby ac faults. Economic power schedules, which reverse power direction, can be made without any restrictions since there is no minimum power or current restrictions.

Urban Infeed. Power supply for large cities depends on local generation and power import capability. Local generation is often older and less efficient than newer units located remotely. Often, however, the older, less-efficient units located near the city center must be dispatched out-of-merit because they must be run for reliable voltage support or inadequate transmission. New transmission into large cities is difficult to site due to ROW and land-use constraints. Compact VSC-based underground transmission circuits can be placed on existing dual-use ROW to bring in power as well as provide voltage support, allowing a more economical power supply without compromising reliability. The receiving terminal acts like a virtual generator delivering power and voltage regulation. Stations are compact and housed mainly indoors making siting in urban areas somewhat easier.

Outlet Transmission for Large-Scale Wind Generation. Large remote wind generation arrays require a collector system, reactive power support, and outlet transmission. Transmission for wind generation must often traverse scenic or environmentally sensitive areas or bodies of water. The VSC-based HVDC transmission allows efficient use of long-distance land or submarine cables and provides reactive support to the wind generation complex.

Multiterminal Systems. The VSC HVDC transmission reverses power through reversal of current direction rather than polarity. This makes it easier to reverse power at an intermediate tap independently of the main power flow direction since voltage polarity reversal is not required. Conventional HVDC transmission requires switching for converter opposite pole connection or polarity reversal.

15.6.3 VSC Station Configuration and Design

HVDC transmission systems based on VSC converter technology are configured as shown in Fig. 15-24.

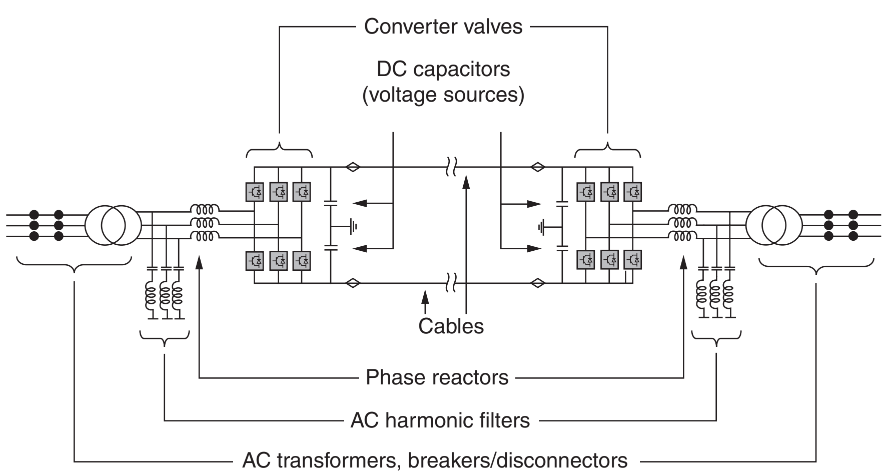


FIGURE 15-24 VSC-based HVDC.

The transmission circuit consists of a bipolar two-wire HVDC system with converters connected pole-to-pole. The dc capacitors are used to provide a dc voltage source. The dc capacitors are grounded at their electrical center point to establish the earth reference potential for the transmission system. There is no earth return operation. The converters are coupled to the ac system through ac phase reactors and power transformers. Harmonic filters are located between the phase reactors and power transformers. Therefore, the transformers are exposed to no dc voltage stresses or harmonics loading allowing use of ordinary power transformers.

A simplified single line diagram for a two-level VSC converter station is shown in Fig. 15-25. Principal station components are described in the following paragraphs.

Power Transformer. The transformer is an ordinary single- or 3-phase power transformer with load tap changer. The secondary voltage, that is, the filter bus voltage, can be controlled with the tap changer to achieve the maximum active and reactive power, both consumption and generation, from the converter. The tap changer is located on the secondary side, which has the largest voltage swing, and also to ensure that the ratio between the line winding and a possible tertiary winding is fixed. The current in the transformer windings contains hardly any harmonics and is not exposed to any dc voltage. In order to maximize the active power transfer, the converter can generate a low frequency zero-sequence voltage ($<$0.2 pu), which is blocked by the ungrounded transformer secondary winding.

The transformer may be provided with a tertiary winding to feed the station auxiliary power system.

Converter Reactors. The converter reactor is installed in series in each phase and is one of the key components in a voltage source converter to permit continuous and independent control of active and reactive power.

The main purposes of the converter reactors are to:

- Provide low-pass filtering of the PWM pattern to give the desired fundamental frequency voltage. The converter generates harmonics related to the switching frequency. The harmonic currents are blocked by the converter reactor and the harmonic content on the ac bus voltage is reduced by an ac filter.
- Provide active and reactive power control. The fundamental frequency voltage across the reactor defines the power flow (both active and reactive) between the ac and dc sides. Refer to typical P-Q diagram and active and reactive power definitions.
- Limit the short-circuit currents.

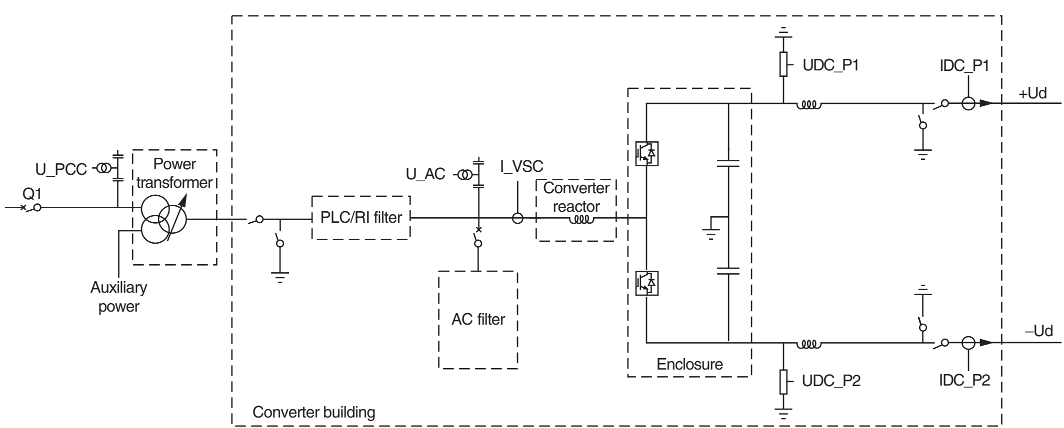


FIGURE 15-25 Simplified SLD for VSC station.

DC-Capacitors. The primary objective of the valve dc side capacitor is to provide a low-inductance path for the turn-off switching currents and provide energy storage. The capacitor also reduces the harmonic ripple on the direct voltage. Disturbances in the system (e.g., ac faults) will cause dc voltage variations. The ability to limit these voltage variations depends on the size of the dc side capacitor. Since the dc capacitors are used indoors, dry capacitors are used.

AC-Filters. Voltage source converters can be operated with different control schemes most of which use pulse width modulation to control the ratio between dc and ac side fundamental frequency voltage. Looking at the ac voltages on the converter side of the reactor, the voltage to ground consists of a square wave as indicated by Fig. 15-3. Connection of a large voltage source converter to a transmission or distribution system requires ac filters to remove the high-frequency components from introducing distortion or interference into the network. This is achieved by means of the converter reactor and the ac filters. The harmonics generated by VSC converters with PWM are higher in frequency than those from conventional HVDC converters. Therefore, smaller filter components can be used to meet performance requirements without large fundamental frequency reactive power generation. This makes the VSC converters better suited to weak-system applications.

The distorted waveform of the converter terminal voltage can be described as a series of harmonic voltages

$$E = \sum_{h=1} E_h \cos(h\Omega_1 t + \alpha_h)$$

where E_h is the hth harmonic EMF. The magnitude of the harmonic EMFs will, naturally, vary with the dc voltage, the switching frequency (or pulse number) of the converter, etc. It will also depend on the chosen PWM control method and topology of the converter. For example, a converter can use sinusoidal PWM with third harmonic injection, that is, when a third harmonic is added on the fundamental frequency modulator to increase the power rating of the converter, or some form of harmonic cancellation such as optimized pulse width modulation, OPWM, can be used. Higher level converters can also be used to switch between a higher number of dc voltage levels, for example, a three level converter can switch between the positive, zero, and negative dc voltage level. In a typical VSC scheme, ac filters contain two- or three-tuned or high-pass filter branches, which can be either grounded or ungrounded.

DC Filters. For VSC converters in combination with extruded dc cables, the filtering on the dc side by the converter dc capacitor and the line smoothing reactor on the dc side is considered to give sufficient suppression of harmonics. However, under certain circumstances, if the dc cable route shares the same right of way or runs close by telephone circuits, railroad signaling wires, or similar, there is a possibility of exposure to harmonic interference from the cable. Under these circumstances and for conditions where a local preventive measure is not feasible, for example, improving the shielding of subscriber wires, the communications company should be consulted for permissible interference limits. A typical requirement can be expressed as an equivalent weighted residual current fed into the cable pair at each station. The current is calculated as

$$I_{eq} = (1/P_{800}) \times \sqrt{\sum_h (P_{hf_1} \times I_h)^2}$$

where I_{eq} = weighted, 800 Hz equivalent disturbing current
I_h = vector sum of harmonic currents in cable pair conductors and screens at harmonic h
P_{hf1} = weighting at the frequency of h times the fundamental frequency

High-Frequency (HF) Filters. In voltage source converters, the necessarily high *dv/dt* in the switching of valves means that the high-frequency (HF) noise generation is significantly higher than for conventional HVDC converters. To prevent this HF noise spreading from the converter to the connected power grids, particular attention is given to the design of the valves, to the shielding of the housings, and to ensuring proper HF grounding connections.

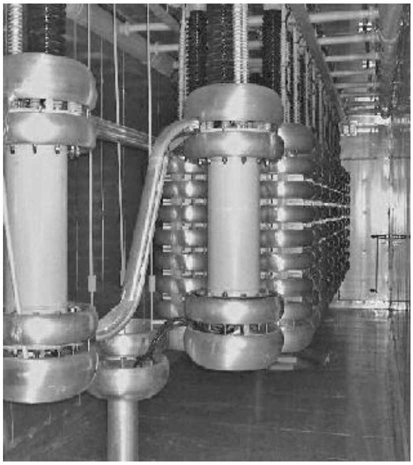


FIGURE 15-26 IGBT valve stacks with corona shields.

IGBT Valves. The insulated gate bipolar transistor (IGBT) valves used in VSC converters are compromised of series-connected IGBT positions. The IGBT is a hybrid device exhibiting the low forward drop of a bipolar transistor as a conducting device (Fig. 15-26). Instead of the regular current-controlled base, the IGBT has a voltage-controlled capacitive gate, as in the MOSFET device.

A complete IGBT position consists of an IGBT, an antiparallel diode, a gate unit, a voltage divider, and a water-cooled heat sink. Each gate unit includes gate-driving circuits, surveillance circuits, and optical interface. The gate-driving electronics control the gate voltage and current at turn-on and turn-off, to achieve optimal turn-on and turn-off processes of the IGBT.

To be able to switch voltages higher than the rated voltage of one IGBT, many positions are connected in series in each valve similar to thyristors in conventional HVDC valves. All IGBTs must turn on and off at exactly the same moment, to achieve an evenly distributed voltage across the valve. Higher currents are handled by paralleling IGBT components or press packs.

15.6.4 Converter Control

The fundamental frequency base apparent power of the converter measured at the filter bus between phase reactor and the ac harmonic filters along with its active and reactive power components are defined by the following equations. Voltage and current phasors used in these equations are according to Fig. 15-27.

$$S_b = P + jQ = \sqrt{3} \times U_F \times I_R^*$$

$$P = \frac{U_F \times U_C \times \sin\delta}{\omega L}$$

$$Q = \frac{U_F \times (U_F - U_C \times \cos\delta)}{\omega L}$$

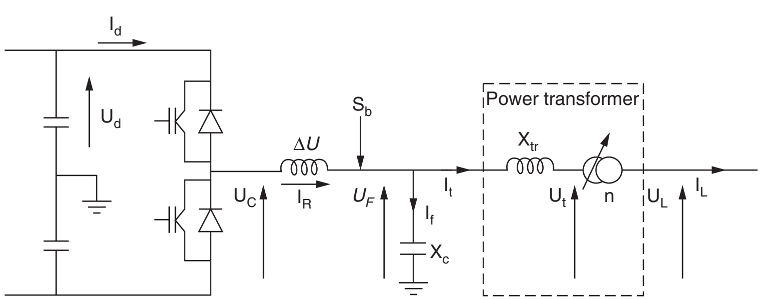


FIGURE 15-27 Voltage source converter.

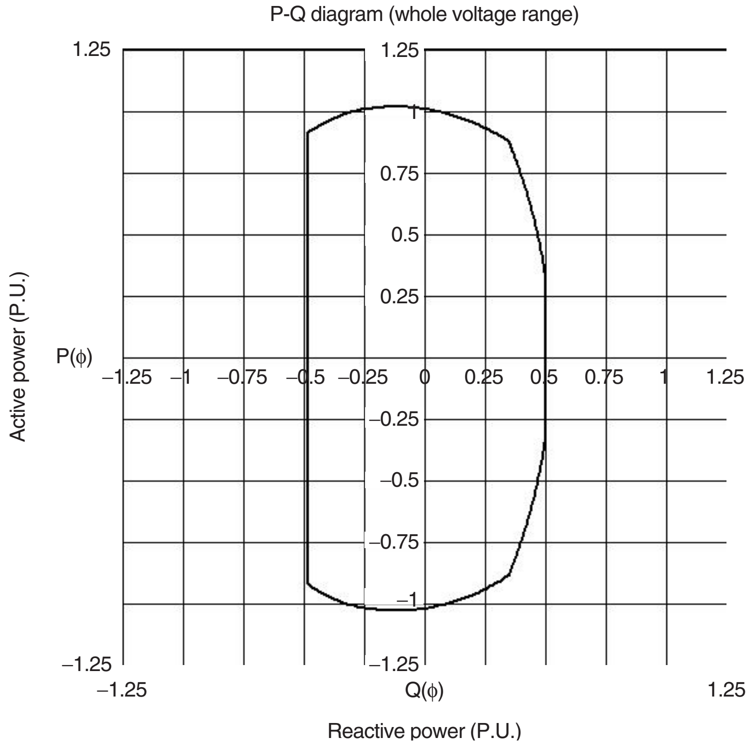


FIGURE 15-28 VSC station net P-Q characteristics with practical limitations.

The inductance of the converter phase reactor is represented by L, and the phase angle between the filter voltage U_F and converter voltage U_C is represented by δ.

The equations illustrate that the power can be controlled by changing the phase angle of the converter voltage with respect to the filter bus voltage, whereas the reactive power can be controlled by changing the magnitude of the converter voltage with respect to the filter bus voltage. By controlling these two aspects of the converter voltage operation in all four quadrants is possible as illustrated in the converter P-Q characteristics shown in Fig. 15-28. This means that the converter can be operated in the middle of its reactive power range near unity power factor to maintain dynamic reactive power reserve for contingency voltage support. It also means that the power transfer can be changed rapidly without altering the reactive power exchange with the ac network or waiting for switching of shunt compensation.

Being able to independently control ac voltage magnitude and phase relative to the system voltage allows use of separate active and reactive power control loops for HVDC system regulation.

The active power control loop can be set to control either the active power or the dc side voltage. In a dc link, one station will then be selected to control the active power while the other must be set to control the dc side voltage. The reactive power control loop can be set to control either the reactive power or the ac side voltage. Either of these two modes can be selected independently at either end of the dc link (Fig. 15-29).

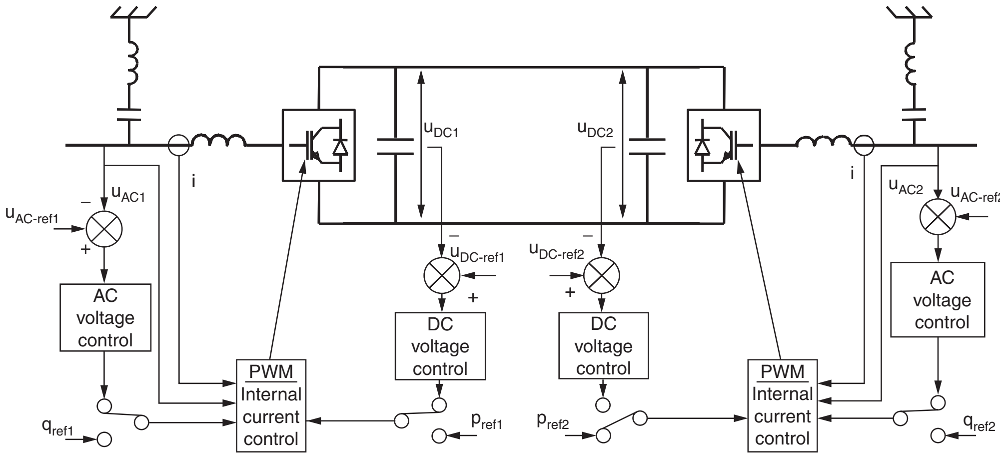


FIGURE 15-29 VSC-based HVDC system control.

15.6.5 Pulse-Width Modulation (PWM) and Harmonic Generation

Pulse width modulation (PWM) of voltage source converters enables independent control of active and reactive power at a constant HVDC voltage using simple two-level converter topology as shown in Fig. 15-30.

A two-level VSC converter can synthesize a balanced set of 3-phase ac converter voltages by injecting either the positive or negative dc voltage on the converter side of the phase reactor. By varying the duration of the positive or negative voltage injections, a sinusoidal voltage with fundamental component at the system frequency can be created. Various PWM switching patterns can be used to minimize harmonics and lower converter switching losses. A PWM pattern with harmonic cancellation or optimized PWM and its harmonic content is shown in Fig. 15-31.

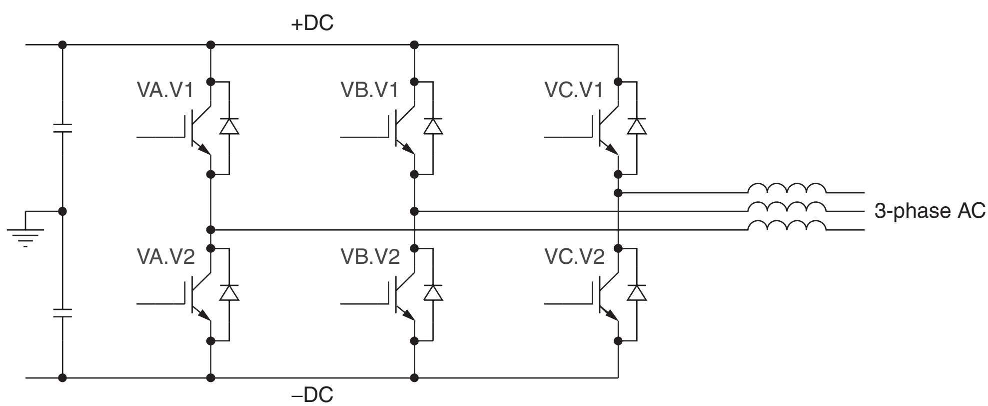


FIGURE 15-30 VSC two-level converter topology.

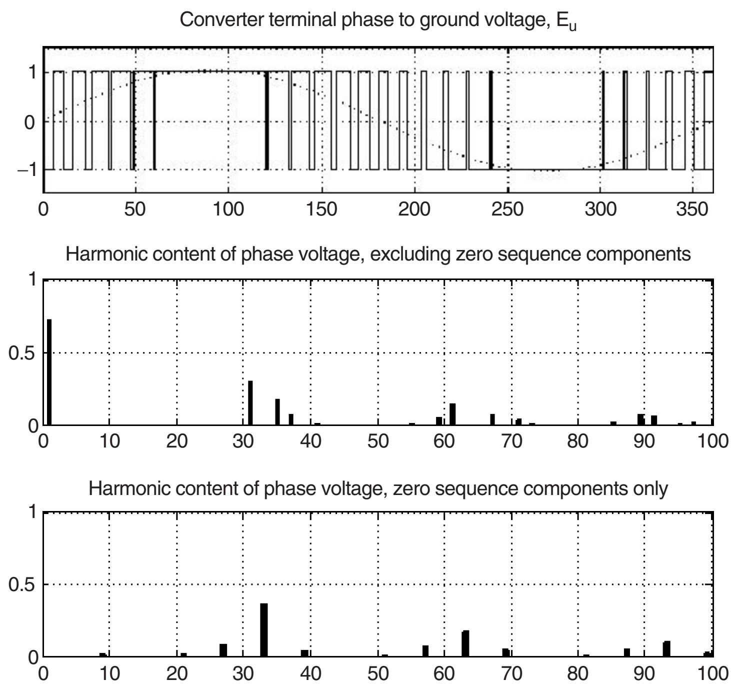


FIGURE 15-31 PWM with harmonic cancellation for two-level VSC.

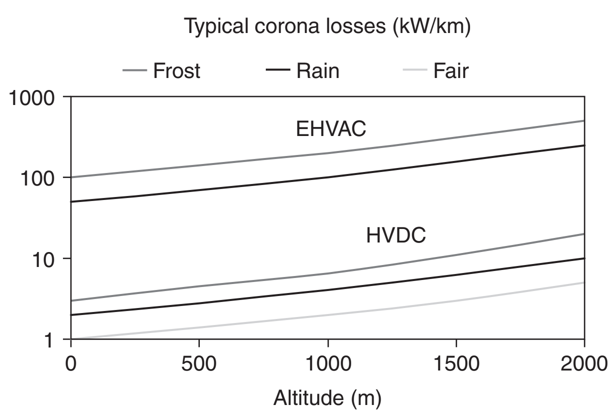


FIGURE 15-32 Foul weather corona loss comparison of EHVAC and HVDC Lines as a function of altitude.

15.7 OVERHEAD LINES AND CABLES

15.7.1 Overhead Transmission Lines

General design criteria for transmission lines can be grouped in the following five categories:

- Power transmission capability
- Power losses
- Insulation coordination
- Corona and field effects
- Mechanical loading

Power transmission capacity is limited by the conductor sag and thermal capacity of the line for the ambient conditions. Emergency loading limits are sometimes used taking into account local conditions and increased sag. These factors affect both EHV ac and HVDC lines. Permissible power transfer levels on EHV ac lines are also affected by surge impedance loading, reactive power compensation, voltage profile, contingency reserve, and stability limits. Transmission on HVDC lines is not limited to reactive power constraints. The HVDC lines cannot become overloaded since the power flow is controlled, therefore contingency reserve is not usually required.

Power losses are due to resistive losses and corona losses. For a given ampacity, resistive losses are lower for an HVDC line than an EHV ac line since the same current is flowing in two sets of conductors in a bipolar dc line compared to three conductors for a 3-phase ac line. Furthermore, the ac resistance is somewhat higher due to skin effect. Although corona losses for EHV ac lines are about the same as those for HVDC lines during fair weather conditions, they increase much more during foul weather conditions, for example, rain, frost, or snow (Fig. 15-32). This means that larger conductor bundles are needed for EHV ac. Dimensions of corona rings are less critical with HVDC. Due to the lower corona levels with HVDC lines, especially during foul weather, fewer bundled conductors are required to meet given requirements on audible noise (AN) or radio interference (RI).

Air clearance requirements are significantly lower for HVDC lines than for EHV ac lines but are more sensitive to altitude effects. Switching surges are significantly lower for HVDC lines than for EHV ac lines. Switching overvoltages govern the clearances for EHV ac lines whereas lightning overvoltages govern the clearances for HVDC lines.

Insulators made of conventional or composite materials can be used for HVDC. The dc operating voltage grading across the insulator string is resistive rather than capacitive. The lower clearance requirements on insulator string length together with the resistive voltage grading make insulator creepage distance more important for HVDC insulators, especially in areas prone to atmospheric pollution. The frequency and intensity of rain are also an important factor since rain washes away accumulated deposits periodically more so on the top surfaces. Additional insulator creepage distance can be achieved with larger sheds, longer skirts, or longer string lengths. A creepage distance of 2.8 cm/kV for lightly contaminated areas can be considered typical. Special considerations exist for insulator cap-an-pin design and choice of materials due to potential for external leakage currents. Collector rings can be used to trap contaminants mitigating uneven deposition along the insulator surface in polluted areas.

There is no electromagnetic induction from HVDC lines. There is an essential difference in acceptance level for dc fields than for ac fields with higher levels for static fields. The International Commission on Nonionizing Radiation Protection (ICNIRP), places a guideline of 40 μT on the maximum static electromagnetic field for continuous exposure to the general public. This compares to the nominal earth magnetic field of 50 μT. The dc magnetic field is very small for two conductors with current flowing in opposite directions at distances several multiples of the conductor spacing.

The HVDC line towers must bear less static and dynamic loading than EHV ac towers due to fewer conductors and insulators. The ROW requirements are narrower with HVDC. In areas where ROW widths are constrained, vertical configurations require less tower height. Balanced structure loading for vertical configurations can be achieved by use of "portal" structures with pole conductors passing through the center of the structure suspended with V-strings.

15.7.2 Underground and Submarine Cables

The HVDC is attractive for higher power transfers over longer distances due to the absence of charging currents and reactive power losses. Fewer cables are needed than for a 3-phase ac circuit. Furthermore, since there is no induction effect with HVDC, cable sheaths do not need to carry the same currents and steel armor can be used for stronger submarine cables.

In ac cables, stress created by the electrical field is distributed in inverse proportion to the capacitance of the cable dielectric. This results in the highest stresses close to the conductor. In dc cables, voltage distribution is determined by insulation resistance and space charges and is dependent on temperature. At higher conductor to sheath temperature gradient, the stress may become highest near the sheath.

Two types of cables are in common use for HVDC transmission, mass-impregnated, nondraining paper-insulated solid cables (MIND), and extruded polymer cables for lower voltage VSC applications (Figs. 15-33 and 15-34). Fig. 15-35 shows voltage waveforms for transformer secondary winding, thyristor valve and dc voltage inside the smoothing reactor. With conventional HVDC, power reversal is achieved by voltage polarity reversal of the cable and 12-pulse harmonic voltages can be imposed on the cable insulation depending on the dc filter design. Fig. 15-36 shows phase reactor voltage, valve voltage and direct voltage for a VSC converter. With VSC transmission, the voltage polarity is constant regardless of transmission direction and switching transients are absorbed by the dc capacitor.

15.7.3 Ground Electrodes

Ground and sea return operation has been used for HVDC transmission to decrease investment costs and lower losses in monopolar submarine cable systems and as a temporary return path for pole outages in bipolar systems. Electrode design always ensures safe step potentials, but other important design factors must be taken into account.

Continuous earth return operation is not always possible due to local soil and geological conditions. With typical earth characteristics, return current penetrates deep within the earth and earth surface potential gradients are low and fall off

FIGURE 15-33 MIND cables for deep sea applications with conventional HVDC.

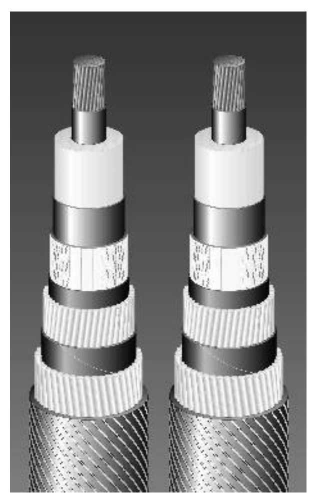


FIGURE 15-34 Extruded polymer cables for deep sea applications with VSC-based HVDC.

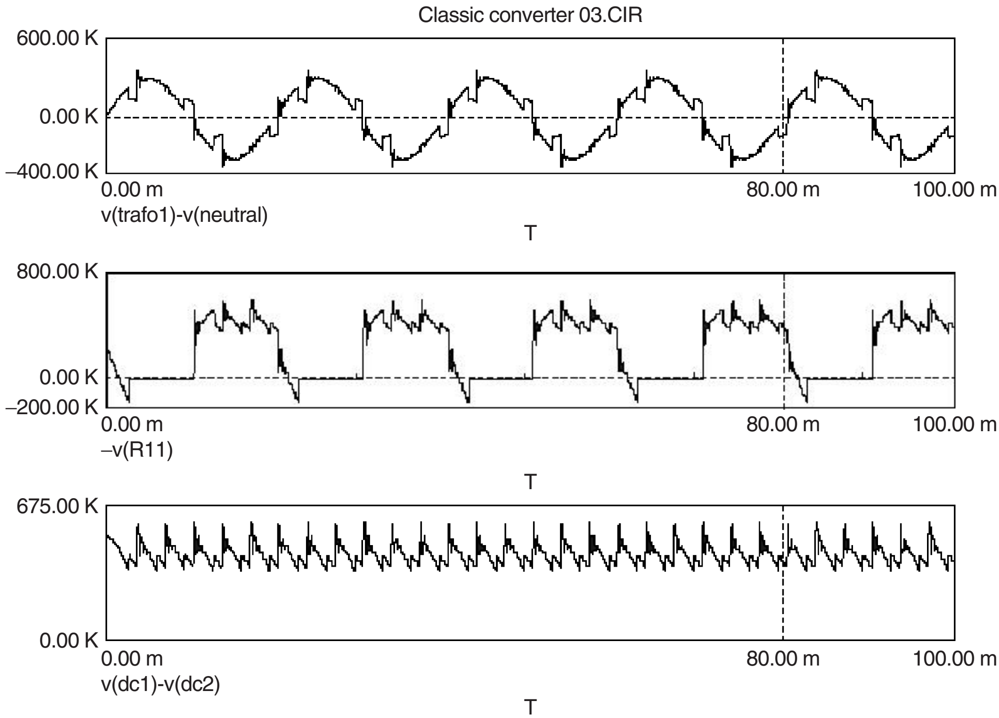


FIGURE 15-35 Voltages for conventional HVDC transmission. Top trace—converter transformer voltage; Middle trace—valve voltage; Bottom trace—direct voltage (inside smoothing reactor).

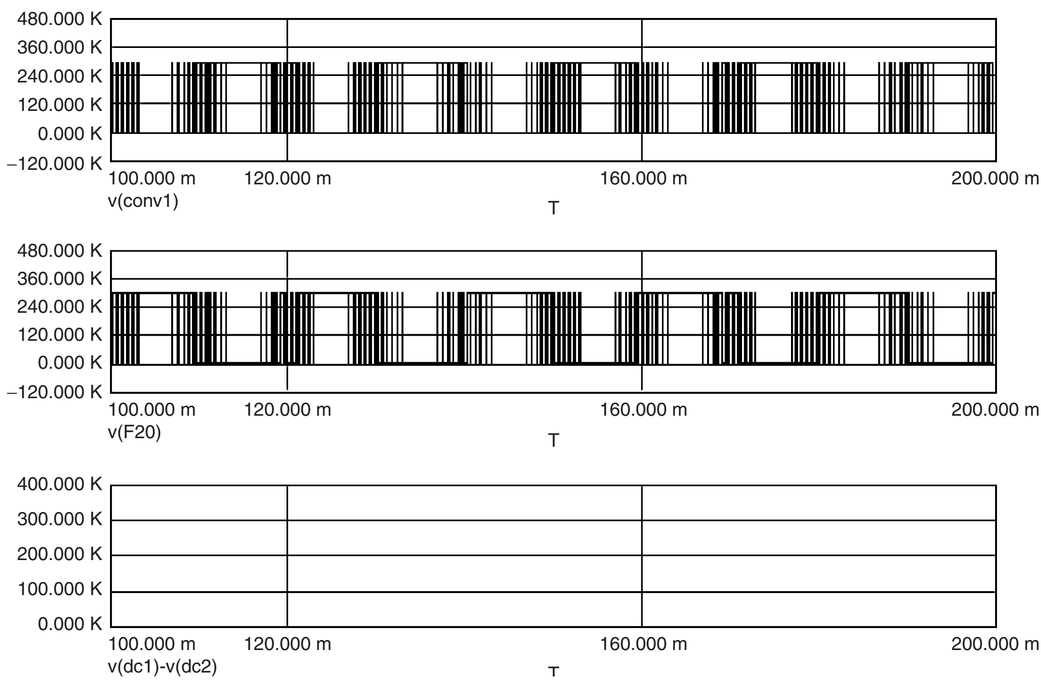


FIGURE 15-36 Voltages for VSC-based HVDC transmission. Top trace—phase reactor voltage; Middle trace—valve voltage; Bottom trace—direct voltage.

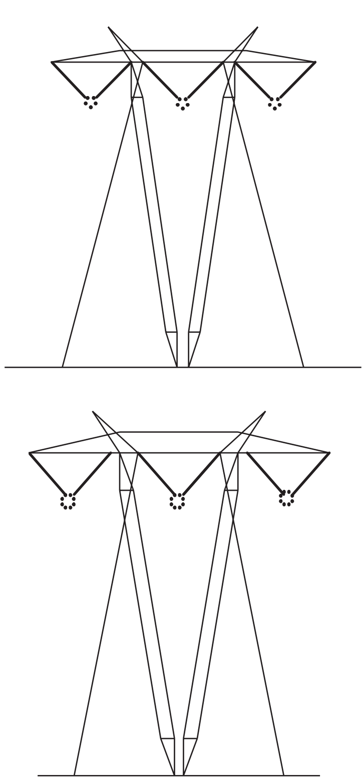


FIGURE 15-37 UHVAC line design.

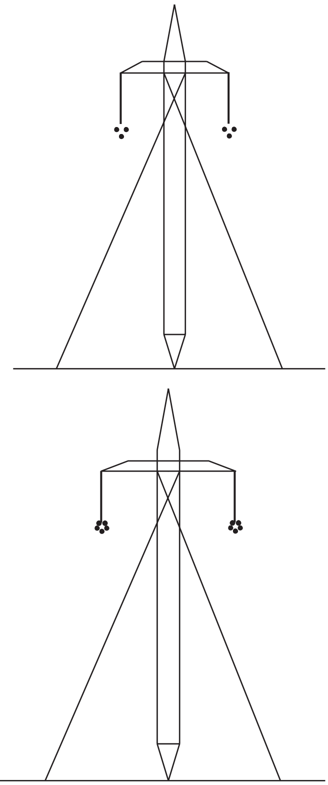


FIGURE 15-38 UHVDC line design.

rapidly with distance from the electrode. In cases with shallow, high-resistivity underlying bedrock, however, the current tends to flow more in the surface layer and the potential gradient extends further from the electrode site. If other conducting underground utilities, such as pipelines, traverse the potential gradient near the electrode, there is risk of stray current pickup and discharge. Over a long period of time, stray current discharge could cause localized corrosion. Corrosion mitigation

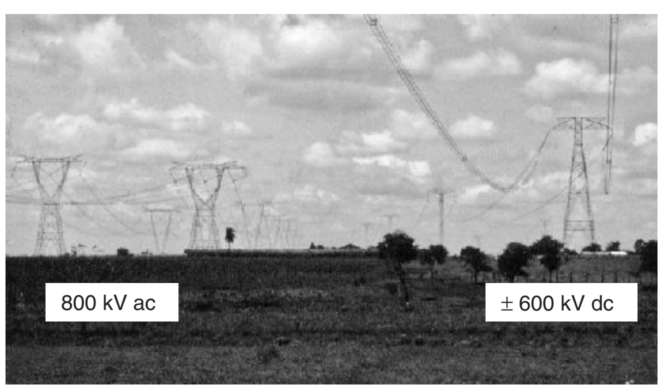


FIGURE 15-39 800 kV EHV ac and ± 600 kV dc.

TABLE 15-3 Comparison of Number of Lines for Given Power Transfer with UHVAC and UHVDC

	kv	Cond. diam. mm	Thermal limit (line) GW	Thermal limit(s/s) GW	SIL GW	1.5 × SIL GW	Required no. of lines 8 GW	12 GW
EHVAC	800	5 × 35	7.5	5.5	2.5	3.8	4	5
	1000	8 × 35	15.0	6.9	4.3	6.5	3	3
HVDC	±600	3 × 50	8.0	5.8	n.a.	n.a.	2	3
	±800	5 × 50	17.7	5.8	n.a.	n.a.	2	3

methods, such as controlled cathodic protection systems, insulating flanges, or sacrificial anodes, can be used or the ampere-hours for earth return operation can be limited through use of metallic return.

15.8 ULTRA-HIGH VOLTAGE DIRECT CURRENT (UHVDC) TRANSMISSION

Most long-distance HVDC transmission systems with power levels above 1000 MW are at a bipolar voltage level of ± 500 kV. Voltage level for the 2 × 3150 MW Itaipu HVDC transmission system in Brazil has been operating at ± 600 kV since the mid-1980s. Transmission voltages of ± 600 kV to ± 800 kV are classified as UHVDC. Higher-power transfers can be achieved over longer distances with lower losses by increasing the dc voltage level into the UHVDC range. A considerable body of work is ongoing in this area for potential applications in China, India, and North America. The controllability and the mechanical and electrical characteristics of UHVDC lines make them in many respects more favorable for long-distance bulk power transmission than UHVAC lines. Figures 15-37 to 15-39 and Table 15-3 compare differences between UHVAC and UHVDC transmission lines.

REFERENCES

Adamson. C. and Hingorani, N.G.: *High Voltage Direct Current Power Transmission*. London, Garraway, 1960.

Kimbark, E.W.: *Direct Current Transmission*. New York, Wiley-Interscience, 1971.

Uhlmann, E.: *Power Transmission by Direct Current*. Berlin, Springer-Verlag, 1975.

Carlsson, L. and Persson, A.: "New Technologies in HVDC Converter Design," *Proc. IEE Sixth International Conference on AC and DC Transmission*, April 29 – May 3, 1996, pp. 387–392.

Jardini, J.A., Reis, L.B., Campos Barros, J.G., and Frontin, S.O.: "HVDC Transmission for Voltages above 600 kV: Evalualtion of DC System and Converter Station Major Equipment," *Proc. IEEE/IREQ International Conf. on DC Power Transmission*, Montreal, 1984, pp. 71–78.

Handbook for *Insulation Coordination of High Voltage DC Converter Stations*, EPRI Report EL5414, Electric Power Research Institute, Palo Alto, CA, 1987.

Fletcher, D.E. and Patterson, N.A.: "The Equivalent Disturbing Current Method for DC Transmission Line Inductive Coordination Studies and DC Filter Performance Specification," *Proc. International Conf. on DC Power Transmission*, Montreal, Quebec, Canada, June 4–8, 1984, pp. 198–204.

Zhang, W., Isalsson, A.J., and Ekstrom, A.: "Analysis on the Control Principle of the Active DC Filter in the Lindome Converter Station of the Konti-Skan HVDC Link," IEEE Paper No. PE-281-PWRS-0-06, 1997.

Kanngiesser, K.W., Bowles, J.P., Ekstrom, A., Reeve, J. and Rumpf, E.: "HVDC Multiterminal Systems," *CIGRE* Paper 14-08, 1974.

McCallum, D., Moreau, G., Primeau, J., Bahrman, M.P., Ekehov, B. and Soulier, D.: "Multiterminal Integration of the Nicolet Converter Station into the Quebec-New Enland Phase II HVDC Transmission System," *CIGRE Proceedings*, Paper 14-103, Paris, 1994.

Bahrman, M.P., Larsen, E.V., Patel, H.S. and Piwko, R.J.: Experience with HVDC-Turbine-Generator Torsional Interaction at Square Butte," *IEEE Trans. Power Appar. Syst.*, Vol. PAS-99, May-June 1980, pp. 966–975.

Lahtinen, M.: "Connection of Harmonic Producing Installations in AC High Voltage Networks with Particular Reference to HVDC: Guide for Limiting Interference Caused by Harmonic Currents with Special Attention for Telecommunications Systems," *Electra (CIGRE)*, No. 159, April 1995, pp. 26–48.

Arkell, C.A., Larsen, K.B., Dellby, B. and Luoini, G: "Comparison of AC and DC Underground and Submarine Cable Transmission Systems," *CIGRE Symposium S 09-87 on AC/DC Transmission Interactions and Comparisons*, Paper No. 500–02, Boston, 1987.

HVDC Ground Electrode Design, EPRI Research Project 1467-1, Report EL-2020, International Engineering Co., Inc.

Asplund, G., Stromberg, H., Blidberg, I. Saksvik, O. and Loof, G.: "Outdoor Thyristor Valve for HVDC," *Proc. IEEE/Royal Institute of Technology Stockholm Power Tech: Power Electronics*, June 18–22, 1995.

Krishnayya, P.C.S.: "Important Characteristics of Thyristors of Valves for HVDC and Static Var Compensators," *CIGRE Conf. Proc.*, Paper 14-10, 1984.

Ekstrom, A. and Liss, G.: "A Refined HVDC Control System," *IEEE Trans. Power Systems*, Vol. 89, 1970, pp. 723–732.

CIGRE Joint Working Group 13/14-08: "Circuit Breakers for Meshed Multiterminal HVDC Systems," Part I, "DC Side Substation Switching Under Normal and Fault Conditions," *Electra (CIGRE)*, No. 163, December 1995, pp. 98-122; Part II, "Switching of Transmission Lines in Meshed HVDC Systems," *Electra (CIGRE)*, No. 164, February 1996, pp. 63–82.

Sakshaug, E.C., Kresge, J.S. and Miske, S.A.: "Arrester Protection of High Voltage DC Transmission System Converter Terminals," Paper No. 71TP47, PWR, Power Appar. Syst., Vol. PAS-90, No. 4, July-August 1971.

CIGRE Working Group 14-05: "Guide for Planning DC Links Terminating at AC System Locations Having Low Short Circuit Capacities," Part I, "AC/DC Interaction Phenomena," *CIGRE Publication* No. 68, 1992.

CIGRE SC 14: 186: "Economic Assessment of HVDC Links," CIGRE Publ. 186.

Asplund, G., Eriksson, K., Jiang, H., Lindberg, J., Palsson, R. and Svensson, K.: "DC Transmission Based on Voltage Source Converters," CIGRE, Paris, 1998.

Axelsson, U., Holm, A., Liljegren, C. Aberg, M., Eriksson, K. and Tollerz, O.: "The Gotland HVDC Light Project—Experience from Trial and Commercial Operation," CIRED Conf., Amsterdam, The Netherlands, June 18–21, 2001.

Railing, B.D., Miller, J.J., Moreau, G., Bard, P., Ronstrom, L., Lindberg, J.: "Cross Sound Cable Project Second Generation VSC Technology for HVDC," CIGRE, Paris, France, 2004.

Weimers, L.: "Bulk Power Transmission at Extra High Voltages. A Comparison between Transmission Lines for HVDC at Voltages above 600 kV DC and 800 kV AC," CEPSI 2004, Shanghai, China, Oct. 17–22, 2004.

Karlsson, T., Liss, G.: "HVDC Transmission with Extremely Long DC Cables, Control Strategies," IEEE/KTH Stockholm Power Tech Conf, 1995, Paper SPT PE 01-05-0638.

Carlsson, L., Flisberg, G.: "Recent Classic HVDC Development," IEEE/PES T&D 2002 Latin American Conf., Sao Paulo, Brazil, March 18–22, 2002.

Holmberg, P., Jonsson, T., Lagerkvist, M: "Properties of Capacitor Commutated Converters in Long HVDC Cable Transmission," EPE'97, Trondheim, Norway.

SECTION 16
POWER-SYSTEM OPERATIONS

Gustavo Brunello
Applications Consultant, General Electric Company

Christa Lorber
Motorola, Inc.

Hesham Shaalan
Associate Professor of Electrical Engineering, U. S. Merchant Marine Academy

Douglas M. Staszesky
Marketing Director, S&C Electric Company

George R. Stoll
President, Utility Telecom Consulting Group, Inc.

CONTENTS

16.1 THE ENERGY MANAGEMENT SYSTEM

16.1.1 Introduction

The management of the real-time operation of an electric power network is a complex task requiring the interaction of human operators, computer systems, communications networks, and real-time data-gathering devices in power plants and substations. There are several concerns that operations departments must take into account in the operation of an electric power system. First and most important is the safety of its personnel and the public. This requires that steps in switching the network be made in accordance with safety procedures so that the lives of utility personnel in the affected substations are not endangered. Next, operating departments are concerned with the security or reliability of the supply of electric energy to customers. In most modern societies, the continuous supply of electric energy is extremely important, and any interruption of a large number of customers at one time is considered an emergency. Finally, the operations department is charged with operating the power system as economically as possible within safety and security limits.

This section deals with the systems that are used to manage a modern utility network. Such a system is usually called an *energy management system* (EMS) and consists of computers, display devices, software, communications channels, and remote terminal units that are connected to control actuators and transducers in substations and power plants. Broadly speaking, these systems are broken down into the following tasks:

Generation control and scheduling

Network analysis

Operator training

The task of managing the generation of a large power system starts with the control of generation to maintain system frequency and tie-line flows while keeping the generators at their economic output. To this are added the economic dispatch, which determines the most economic output of each generator for a given load, the on/off scheduling or commitment of generators to meet varying load demands, and the determination of the pricing and amount of energy to buy and sell with neighboring utilities.

The task of managing the transmission system network requires the monitoring of thousands of telemetered values, the estimation of the electrical state of the network given the telemetered values, and the estimation of the effect of any plausible outage on the operation of the network. The security-analysis problem requires that the EMS be capable of analyzing hundreds or thousands of possible outage events and informing the operator of the best strategy to handle these outages if they result in an overload or voltage limit violation.

The operators must be highly trained in the use of the EMS and how to respond to emergencies. To be sure that operators are trained effectively, most utilities incorporate a simulator into their EMS that is capable of simulating the effects of an emergency on the power system. The operator is then required to "respond" by taking actions on the simulator that corrects the emergency problem. In this way new operators can be introduced to emergency procedures and experienced operators can have their training refreshed.

The EMS systems now in use in a modern power-system operations department are very large computer systems that require a large maintenance staff. The EMS is usually one of the largest computer systems in use in a utility company and often has within its database the needed information for many of the other engineering and design departments. In recent years, the concept of open systems has taken hold within utility EMS systems so that they are approaching a truly distributed form of command and control system.

16.1.2 Overview of Energy Management System Functions

Supervisory Control and Data Acquisition (SCADA) Subsystem. Supervisory control supports operator control of remote (or local) equipment, such as opening or closing a breaker, with security features, such as authorization and a select-verify-execute procedure. The data-acquisition subsystem gathers telemetered data for use by all other functions within the EMS. Data are obtained from various sources including remote terminal units (RTUs) installed in plants and substations and devices near to the system control center by local input-output (I/O) equipment.

A SCADA system provides three critical functions in the operation of an electric utility network:

Data acquisition

Supervisory control

Alarm display and control

Data-Acquisition Function. The data-acquisition subsystem periodically collects data in processed or raw form from remote terminal units. Data acquisition consists of five functional areas:

Data collection

Data processing

Data monitoring

Special calculations

Scan configuration control

Data collection is responsible for periodically acquiring data from remote terminal units at the appropriate rate. In addition, data collection monitors the various scans to make sure they initiate and complete within the current time period.

Data processing is responsible for converting analog values from raw data to engineering units. It is also responsible for converting digital status points to a system convention of device states (0 for closed and 1 for open). Data for points that are manually replaced in the database are not usually processed. Data processing is also responsible for handling data obtained from data links to other computer systems.

Data monitoring interfaces with the alarm processor and notifies it when the following occur:

Devices change state

Values exceed operating limits

Data monitoring also provides deadband and return-to-normal features.

Special calculations support various standard calculations such as

Copy a value

MVA from MW and Mvar measurements

MVA from kV and amperes

Amperes from MVA and kV measurements

Other common periodic calculations

Calculated values are derived periodically from scanned data in the database.

Scan configuration control removes a terminal unit from the scan or switches the channel assignment when sustained communications errors occur. Scan configuration control periodically attempts to reestablish communications with terminals, which have been removed from the scan.

Supervisory Control Function. This function allows the operator to control remote devices and to condition or replace values in the database. All operations are multistep procedures. Selection of the device to be operated is the first step. Next is the visual verification step, and the final step is operator execution or cancellation. Data conditioning includes operations such as the following:

Manual replacement of telemetered data

Alarm inhibit/enable

Reverse normal (change definition of the normal state of a device)

Bypass enter (of failed telemetry)

Tag/tag clear

Summary displays support the manual replace, alarm inhibit/enable, and tag/tag clear functions. Entries on these summaries are typically in inverse chronological order, the most recent entry being at the top of the summary.

Alarm Display and Control Function. The subsystem is responsible for the presentation of alarms to the operator. It supports alarm presentation and alarm presentation control. Alarm presentation is responsible for constructing the alarm message, organizing alarms in categories, maintaining an alarm summary display and abnormal summary, maintaining console logs, initiating audio/visual annunciators, and interfacing to other functions (e.g., the mapboard). Presentation control assigns priorities to alarm messages, recognizes points which are inhibited from alarming or manually replaced by the operator, and provides operator functions such as alarm acknowledgment.

User Interface Subsystem. The most visible feature of an energy management system is the user interface (UI) subsystem, which includes the following:

Presentation of system data on visual displays

Entry of data into the EMS through a keyboard

Validation of data entry

Support of supervisory control procedures

Output of displays to a printer or video copier

Operator execution control of application programs

Displays are created by using an interactive display builder, which allows definition of linkages between areas on the display and the EMS database for retrieval and entry of data. Also, the user can define function keys or function keys/display locations (poke points) when building a display to cause the presentation of another display or to initiate the execution of an application program.

The display builder allows the operator to create or modify the static elements of the display and add, modify, or delete the data and control linkages of the display. When the operator is satisfied with the display, the display definition is saved in the display file for later use by UI.

Displays are presented on a cathode ray tube (CRT) display at a console. An EMS console consists of one or more CRTs having full graphics capability, a display controller, a keyboard, and a trackball or mouse.

The flexibility in display format provided to the user allows a single subsystem to support a wide range of display types. These typically include

Menu or index displays

One-line schematic circuit diagrams

System overviews
Substation and generation displays
Transmission line displays
Summary displays
System configuration displays
Application program displays
Trend or plot displays
Disturbance data collection displays
Historical data storage displays
Report displays
Other displays

Communications Subsystem. The communications subsystem encompasses management of a local-area network supporting the EMS itself, such as a dual-redundant Ethernet, token ring, or fiberoptic communications medium, and support of communication with other computing systems and field equipment.

In addition to the users within the control room, there may be schedulers, trainees, programmers, engineers, and executives who require access to the EMS through standard console displays, remote displays, or even personal computers. All these have to be connected to the EMS via a local area network that may extend outside the control center building to other facilities.

Other connections within the utility may include off-line engineering systems for planning or long-range scheduling, other control systems, for example, load management, distribution, or plant management, and control and corporate (billing and customer) computer systems. External communications are typically with other utilities or power pools.

Information Management Subsystem. The information management subsystem supports definition of and access to data used by the EMS. This includes all the static data descriptive of the power system, the EMS configuration, and data shared with other systems. It also includes organization of data for specific uses, for example, for data acquisition and monitoring and for network analysis algorithms.

In current EMS configurations, the database is distributed. This results in a need to facilitate data access without burdening either the operator or the applications programmers and other system users. Evolution of software standards and tools in the computer industry has led to products that support these needs, such as relational database managers and computer network file and resource managers.

Applications Subsystem. The applications extend the usefulness of an EMS, allowing data gathered by the SCADA system to be used to optimize and control the power system. An EMS overview is shown in Fig. 16-1.

Generation Control Applications. An interconnected system is made up of one or more control areas, each of which is defined as that portion of an interconnected system to which a common generation control scheme is applied. It also may be regarded as that portion of the interconnected system which is expected to regulate its own generation to follow its own load changes. It may consist of a single utility, or a part of one, or a whole group of pooled utilities. In each case, a control area would include all the generating units, loads, and lines that fall within its prescribed boundaries. All the control areas of an interconnection, taken together, should account for all the generation, load, and ties of the interconnected system.

A single-area system is one in which the entire interconnected system is encompassed within one control area. One control system provides the basic regulation for the entire interconnection and does not distinguish between the locations of load changes within the interconnection. A multiple-area system is one in which there are many control areas, each with its own control system, each normally

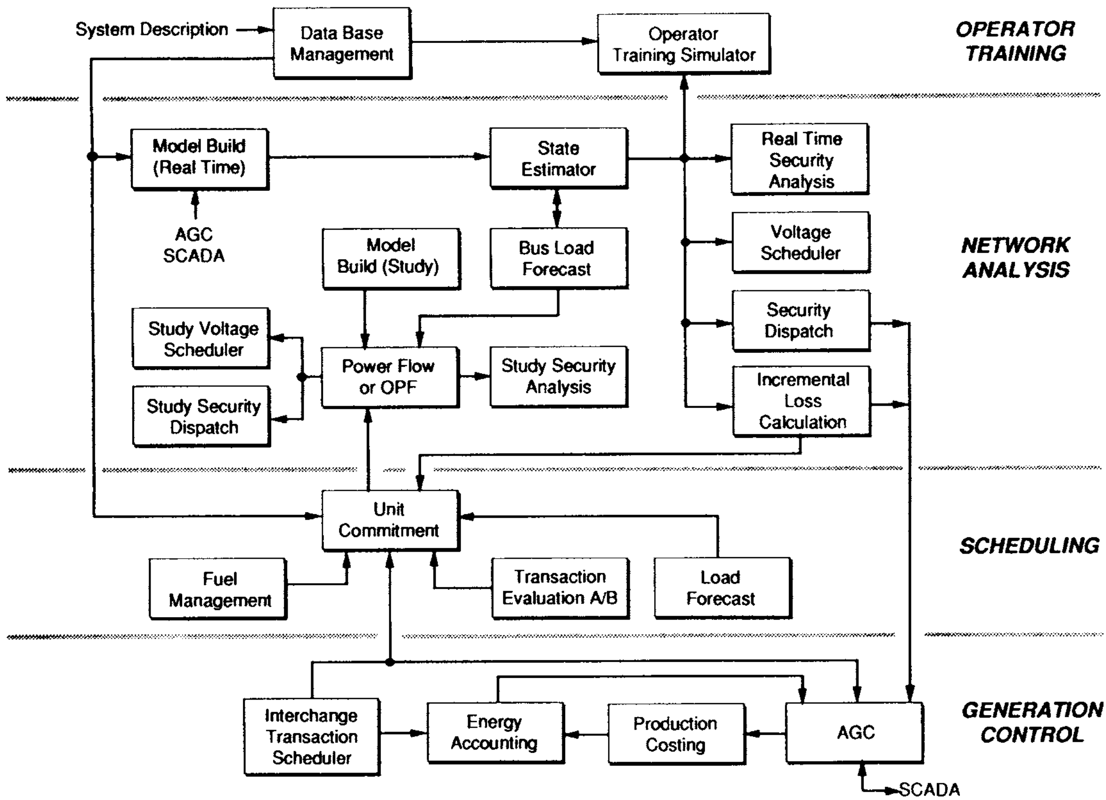


FIGURE 16-1 Energy management system.

adjusting its own generation in response to load changes within its own area. All the interconnected systems in the United States and Canada operate on a multiple-area basis.

Speed Governor. The generating unit's speed governor, along with governor-controlled steam valves (in a thermal plant) and a speed changer which provides for adjustment of the governor set point, constitutes the primary control loop for maintaining frequency at the unit level. The steady-state speed regulation characteristic of the speed governor relates a per-unit change in rated speed (y axis) to a per-unit change in rated load (x axis) and is a straight line with negative slope (called *droop*). Thus, with the speed changer set to provide rated speed for a given load, changing the set point shifts the straight-line characteristic along the x axis so that more or less output is demanded for constant rated speed. The automatic generation control (AGC) signal to raise/lower the set point (or signal for a directed set point) closes the system-level control loop and is also referred to as *supplementary control.*

Operating Objectives of Generation and Power-Flow Control. Automatic control of generation and power flow is an essential need for the smooth, neighborly, and effective operation of a widespread interconnected system. On a multiple-area interconnection, the regulating or control objectives are threefold:

Objective 1. Total generation of the interconnection as a whole must be matched, moment to moment, to the total prevailing customer demand. This in itself is achieved by the self-regulating forces of the system.

Objective 2. Total generation of the interconnected system is to be allocated among the participating control areas so that each area follows its own load changes and maintains scheduled power flows over its interties with neighboring areas. This objective is achieved by area regulation.

Objective 3. Within each control area, its share of total system generation is to be allocated among available area generating sources for optimum area economy, consistent with area security and environmental considerations. This objective is achieved by economic dispatch, supplemented as required by security and environmental dispatch.

The means of achieving objectives 2 and 3 are referred to as *supplementary control,* or currently—and more generally—as AGC. Such control may be regarded as a reallocation control redistributing the systemwide governing responses to load changes in various areas to generators within the areas that had the change. Each area then follows its own load change, with scheduled internal distribution. On a single-area system, objective 2 does not apply.

These functions act at the overall system level to regulate the real power output of generation, economically allocate demand among committed units, calculate various reserve quantities, determine production costs, and account for interchange of power between utilities and/or control areas.

Automatic Generation Control. Automatic generation control, sometimes called *load-frequency control* (LFC), regulates power system in terms of maintaining scheduled system frequency and scheduled net interchange. Automatic generation control is implemented as a closed-loop feedback controller. The error signal is determined either as a computed area control error (ACE) for a control area or a given area requirement (AR) in some power pool control structures. Positive ACE indicates overgeneration; positive AR indicates undergeneration. The ACE calculation is based on frequency deviation from schedule, net interchange deviation, or a composite tie-line bias. In tie-line bias control mode, interconnected control areas jointly participate in maintaining frequency, which is uniform among areas, but are individually responsible for maintaining each area's scheduled net interchange. The formula for this is

$$\text{ACE} = B(f_{\text{actual}} - f_{\text{scheduled}}) + (\sum \text{TMW} - I_{\text{scheduled}})$$

where the summation is over all tie-line megawatts (TMW), I is the current scheduled net interchange level, and B is tie-line bias, which converts frequency deviation to real power, usually expressed as MW/tenth Hz. B is characteristic of the installed capacity (MW) of the control area and is usually a constant. Additional terms or modifications to the formula are used to account for correction of time errors, inadvertent interchange payback, and so on.

Area control error is a noisy signal and so requires processing. Processing also includes provision for proportional, integral, and anticipatory (or derivative) control characteristics for AGC as a feedback controller. Integral control is necessary to prevent long-term offset in frequency and to ensure that ACE crosses zero (the normal set point) frequently. System control requirements thus determined from processed ACE are allocated to generating units based on several criteria.

Unit Control Considerations. Key considerations are

The deviation in each unit's loading from the most recent economic assignment—MW level

The deviation of total system load since the last economic dispatch

The current value of ACE

Economic base points are assigned by the economic dispatch (ED) function, and LFC will drive unit loading toward these assignments unless there are overriding conditions. This mode is termed *mandatory* unit control (mandatory with respect to economics).

An overriding condition may be that ACE exceeds a threshold beyond which correcting ACE takes precedence. In this case, AGC is operating in a *permissive* mode (with respect to economics). Here units are inhibited from moving against correction of ACE. If ACE exceeds a larger threshold, an *emergency assist* mode is entered. Here all units move to correct ACE and may move against their economic directions, that is, away from economically assigned base points.

Units participate in ACE reduction in proportion to regulating participation factors, which may be operator-entered or calculated from various criteria according to individual company or pool operating policies. Units participate in adjusting to the deviation in system load since the last ED by use of economic participation factors, produced by ED. In some systems, a single set of participation factors is used.

Unit desired generation is calculated according to the preceding rules, and control output is sent to generating station RTUs either as MW set points or raise/lower signals as appropriate to the local generating unit plant-control equipment.

Control of each unit assigned to automatic regulation is performed by a separate unit-control loop (feedback controller). Here the set point is unit desired generation already obtained. Models of individual unit dynamic response to previously issued control commands are compared with actual telemetered output of the unit in determining the degree of new control to be issued.

AGC Operator/Dispatcher User Interface. Typical AGC displays used by system operators include

System summary—provides an overview of system control information such as area control error, reserve quantities, incremental costs, lambda (from ED), and AGC control mode states and allows the operator to change these states or enter key parameters.

Generation summary—summarizes current status and output of all generating units and may provide for operator changes to unit status.

Station/plant summary—shows detail related to operation of individual units, limits, fuels, costs, and so on.

Tie-line summary—shows telemetered real and reactive power flow on all tie lines and net total real power interchange and may show line limits.

Figure 16-2 shows an overview of a typical AGC program.

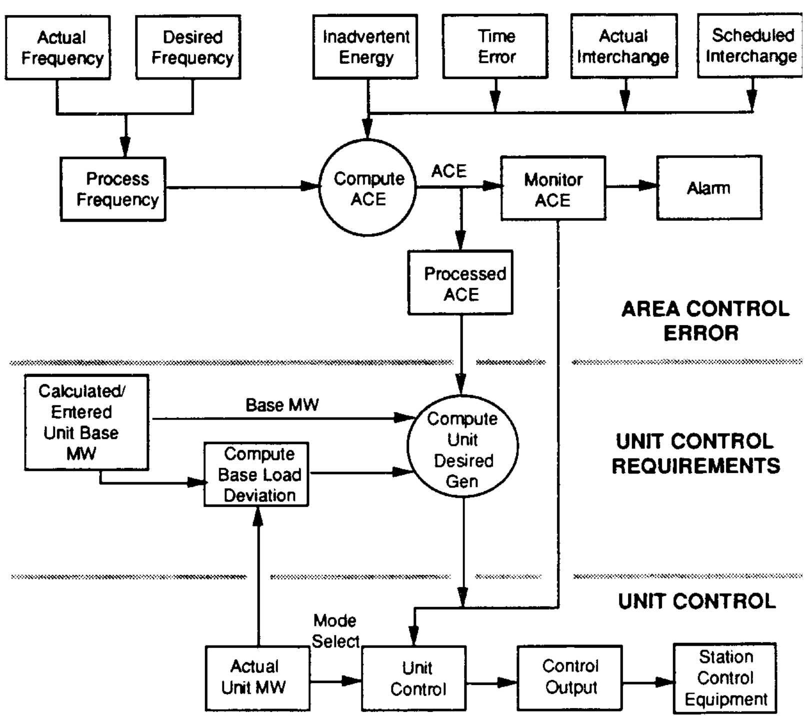


FIGURE 16-2 Overview of an automatic generation control system.

Interchange Scheduling. The interchange transaction scheduler (ITS) function supports the operator in entering (defining), editing, and reviewing power interchange schedules with neighboring control areas/utilities. The schedules are usually negotiated by the operator over the telephone with other operators in control rooms at other utilities. These schedules are utilized principally by AGC and energy accounting.

Schedules are established by utility and by account within each utility. Examples of accounts include firm or nonfirm energy and capacity purchases, sales, and so on. Schedules may be defined on a daily hour-by-hour basis or on a start/stop date and time basis according to company or pool operating procedures. Various entry displays support definition of such schedules. Other displays are used to summarize transactions by company, account, or chronology.

Given a multitude of concurrently active transactions, a net profile of interchange is constructed in order to provide AGC with the instantaneous net scheduled interchange needed for real-time system regulation. At the end of each hour, scheduled transactions are compared with actual data in the energy accounting function to maintain historical records.

An emergency scheduling capability allows the operator to enter a single net schedule of interchange to override all other currently active schedules. Other entries associated with transactions may include cost, price, ramp rates (MW/minute), and additional information associated with third-party or "wheeling" transactions.

Economy A Transaction Evaluation. Economy A Transaction Evaluation is a user-oriented program for evaluating short-term interchange transactions with a neighboring utility. It applies to transactions, which do not involve altering the commitment of generating units.

The idea behind Economy A transactions is to find an amount of power to interchange with a neighboring system so that both systems achieve maximum benefit. Essentially, this means that the system with lower incremental cost of generation will sell power to a neighbor with higher incremental cost. The optimal amount of power interchange is that which brings the two systems to the same incremental cost.

To find the optimal interchange, agreed increments or blocks of interchange are added or subtracted to the base economic dispatch. For each block, a price or cost increment is calculated. The operators in each system then use the block information to determine the number of blocks to use in reaching a final interchange value.

The program also can use the economic dispatch package in a study mode to calculate incremental and production costs under a variety of conditions specified by the operator. Parameters for these calculations can include generation conditions, interchange schedules, and unit costs.

Input. Economy A obtains the following from automatic generation control:

Economic and operating limits, mode, and assigned or base generation

Fuel costs

Starting megawatts

Efficiency factor

Heat-rate curve selection

Operator inputs consist of requests, modification of the preceding data, and definition of the transaction and system parameters.

Output. Results of Economy A Transaction Evaluation are presented in CRT displays and also can be sent to a printer. This output includes

System results, such as production costs, spinning reserve, and incremental losses, for each block evaluated

Economically assigned generation for each unit

Energy Accounting. The energy accounting (EA) function maintains accumulated operating data in accounts ordered on an hourly, daily, monthly, and/or yearly basis. These accounts typically relate to energy exchanged via tie lines, plant generation, large-customer consumption, and on/off peak cumulative inadvertent energy exchanges. Additional data such as production costs or purchase/sale costs also may be accumulated, and in a hydroelectric system, discharge of water or pond levels may be recorded. In practice, generalized calculation and report functions are configured to provide energy accounting capabilities.

Accumulating energy data is accomplished either by field equipment such as pulse accumulators (counters) which provide energy data to be telemetered or by telemetering power (megawatt) values to the EMS, where these are integrated to obtain energy data (MW-hours).

Daily power system values are collected on an hourly basis. Correspondingly, monthly values are collected and stored once a day so that there is a value for each day of the month. The following paragraphs describe typical energy accounting processing that is performed on either a daily or monthly basis.

Daily Features. Energy accounting collects the instantaneous tie-line megawatt values every minute and at the end of the hour produces the integrated values for all tie lines. It then subtracts these values from the corresponding tie-line pulse accumulator values and stores the difference. The absolute difference is compared with a tolerance (for each tie line). This allows the accuracy of tie-line telemetry information to be continuously monitored.

Energy accounting maintains actual tie-line data for each hour of the day. It also classifies the values according to whether the hour of the day is an off-peak or on-peak hour. On-peak and off-peak start and stop times are defined via the information management function. Holidays and Sundays are considered off-peak. This allows interchange (both actual and scheduled) and inadvertent calculation to be divided into on-peak and off-peak accumulations. Daylight savings time conversion days (23-h or 25-h days) are also supported. For these days, the appropriate amount of data is collected and processed accordingly.

At the end of each hour, the hourly actual interchange values collected are added into running totals of on-peak and off-peak energy (depending on the hour). The scheduled interchange values provided by ITS are also added to on-peak and off-peak accumulations. Following the accumulation of interchange (scheduled and actual), the inadvertent energy for the hour is computed as the deviation between actual and scheduled interchange.

The inadvertent energy value for the hour is then saved. The hourly value is then used to update the cumulative (on-peak or off-peak) inadvertent energy value. The appropriate cumulative inadvertent energy value is then made available to AGC.

Energy accounting also may collect and maintain production cost data for each hour of the day. At the end of each hour, the production cost data for each generator and the system are collected and stored. Additionally, energy accounting supports the calculation and storage of system net generation and control area net load for each hour of the day. For all values maintained on a daily basis, the running daily total for each quantity is also updated and retained.

Production-Cost Calculation. Production costing (PC) calculates the hourly production cost for each generating unit and the entire system. Production costing is synchronized with execution of the economic dispatch program and supports the following features:

Production costing executes periodically throughout the hour, and the average hourly production cost is calculated at the end of the hour.

Several sets of production cost values can be calculated from the current actual unit generation levels and for the generation levels recommended by the economic dispatch.

System dispatch performance is monitored by computing actual generation costs, dispatched production costs, and ideally dispatched production costs (manual dispatch).

A set of unit fuel consumption values can be computed from actual unit generation values.

Unit and system daily logs are provided showing all relevant hourly and daily values via the energy accounting and reporting support functions.

The periodic production costs are calculated by integration of the area under the incremental cost curves or by separate I/O curves and can include the effect of incremental and fixed maintenance costs, fuel cost, and efficiency.

The periodic unit actual fuel consumption is calculated and includes the effect of the unit's efficiency. The unit actual fuel consumptions are summed to yield the current system fuel consumption. All unit production costs are summed to give the system production cost values.

The periodic values are integrated over the hour to produce hourly unit fuel consumption and production cost values. The hourly production costs and fuel consumption values are saved at the end of each hour. These values are then stored in a historical database by energy accounting.

Generation Scheduling Applications. The forecast and scheduling applications within an energy management system gather, organize, and use large amounts of historic and economic information. This group of related software packages puts that information to work in forecasting loads, scheduling units and generation, evaluating Economy B type transactions with other utilities, and tracking fuel contracts. Forecast and scheduling applications are tailored to the power system they serve. For example, a unique load forecast model is developed for each case.

Load Forecast. This program forecasts hourly loads 1 to 7 days in advance. Load-forecasting methods are based on similar days according to season, day of the week, and so on, with further adjustment for weather effects by using

Nonlinear, dynamic, adaptive weather model
Correlation of load to temperature, humidity, light intensity, and wind speed
Adaptation to real-time load and actual weather conditions

Unit Commitment. This program schedules hourly status (on line/off line) and output for each online unit, 1 to 7 days in advance. The calculations consider

Production cost models
Start-up cost model
Shutdown cost
No-load (spin) cost
Incremental maintenance costs
Network losses

Unit commitment runs with two sets of constraints. System constraints are

Load forecast
Interchange schedules
Reserve requirements
Regulation requirements

Unit constraints are

Prescheduled status or output
Derations
Multiple limits
Rate limits
Up- and downtime limits
Reserve limits
Plant start-up limits

Each unit can be assigned these models

Thermal
Combustion turbines
Combined cycle
Ramping times
Start times
Multifuel

Economy B Transaction Evaluation. Economy B transactions are similar to Economy A except that generating units must be added or taken off line to meet the contract. This program does a before-the-fact evaluation of proposed interchange transactions. After the fact, it can make the same analysis to evaluate the worth of each transaction. It can

Perform multiple commitments against levels of prioritized interchange
Recommend prices
Make buy/sell analysis
Use fixed, operator-entered, or variable prices

Fuel Management. The fuel-management programs incorporate fuel constraints into unit commitment schedules so as to optimize the use of fuel contracts. Contracts can be

Take or pay
Fixed price
One hour to one month

Contract limits can be

Hourly to monthly
Rate of consumption
Total consumption

Network Analysis Applications. These monitor the security of the system and assist the operator in optimizing system performance. The *model-build program* responds to switching operations in the transmission system. With this information it determines the current network configuration. This constantly updated real-time model is used by other network analysis programs.

Inputs to the program are all measurements (including MW, Mvar, kV, and amperes), zero injections, and calculated loads. The *state estimator* uses statistical methods to check for bad data and to establish a consistent network solution as a basis for security analysis and power flow studies.

The *bus-load forecast* provides a forecast for each individual bus, for any specified hour of the week. Forecasts are based on the history of user-defined load groups. Both MW and reactive ratio histories are used. This information is used for studies and also can be used to support temporarily outaged telemetry.

Voltage scheduling is an optimization program that minimizes power losses in the system by adjusting unit voltages, load tap changing (LTC) taps, and phase-shifter taps. The program performs this optimization while maintaining voltages and Mvars within permissible ranges.

Optimal power flow (OPF) enables the operator to study a network solution, which describes the steady-state power flow that would result from specified network conditions. It can optimize system variables to enhance power system security and/or economy.

Security analysis determines the security of the power system under specified contingencies. It stimulates the steady-state power flow for each case and then checks for out-of-range conditions. Security analysis also handles split bus, altered topology, and islanded systems.

Security dispatch detects overloads in the real-time network model and determines control actions such as generator shifts that will alleviate the overload or that will avoid an overload after a contingency. The program can incorporate phase shifters, interchanges, and load shedding as well as unit outputs to solve problems.

Operator Training Simulator. With an operator training simulator (OTS), it has become possible to improve the quality of training for power system operators. The OTS allows operators to be exposed to simulated power system emergencies and to practice alleviating these emergencies. Similarly, operators may practice system restoration under simulated conditions. Since operators may be exposed to simulated emergency and restorative conditions on the OTS, frequently and at will, as opposed to rarely and by chance on the job, the time required to train a new operator may be significantly shortened. Similarly, with an OTS, it becomes possible to expose experienced operators to emergencies and restoration procedures as part of refresher training.

The simulator can present results to the operators, which are as accurate as those observed by the EMS using typical power-system telemetry. The operator uses a user interface and applications functions which are identical in the OTS and in the EMS.

The OTS includes long-term dynamic models of the electrical network, loads, generators, turbines, and boilers. The OTS also includes the control functions of the EMS: SCADA, power applications, and their user interface. In addition, an educational subsystem is provided with features that allow the instructor to construct groups of one or more training events or power system disturbances and to store and retrieve these groups of events.

Other significant features of the OTS include

The power-system model in the OTS is the same as the model used in the EMS.

The OTS uses multiple consoles to support team training and an instructor position.

The OTS supports a load model which includes the effect of frequency, voltage, load management, and subtransmission reactive shunts and taps.

The OTS supports system restoration/blackstart exercises.

Underfrequency load shedding is modeled in the OTS.

The OTS allows representation of a wide range of power-system events or disturbances.

The OTS may include a model of the AGC systems of external companies.

The OTS includes relay models for over/undervoltage, inverse time overcurrent, over/underfrequency relays, synchro check relays, time switching, volts/Hz, over/underexcitation, and automatic reclosure.

The OTS includes features that allow the instructor to play the role of power-plant operators, substation operators, and neighboring company operators.

OTS Functional Description. The overall simulator system can be logically divided into four principal subsystems: the power-system model (PSM), the control-center model (CCM), the educational system, and the user interface.

The PSM simulates response of load, generation, and network conditions (flows and voltages) to control actions, which were initiated either by the operator or by AGC, and to preset events from the training system. The PSM includes a load-model program, network modeling, which is implemented as a network topology processor, and a fast decoupled load-flow algorithm and a set of prime mover models and frequency-response programs. The control-center model includes a replica of the control functions in the EMS. Included are the SCADA/AGC functions and selected network analysis functions. The educational subsystem provides a means for sequences of events to be defined, stored, and retrieved by the instructor. Separate displays are used to define each sequence and to catalog by title those presently stored. The user interface relates to all the previous subsystems. It provides display and control, via the workstation display and keyboard, and logging of all system events.

The operator simulation process differs from the operating models primarily in the time frame considered. Transient time scales are on the order of cycles (0.016 s), and longer dynamic stability

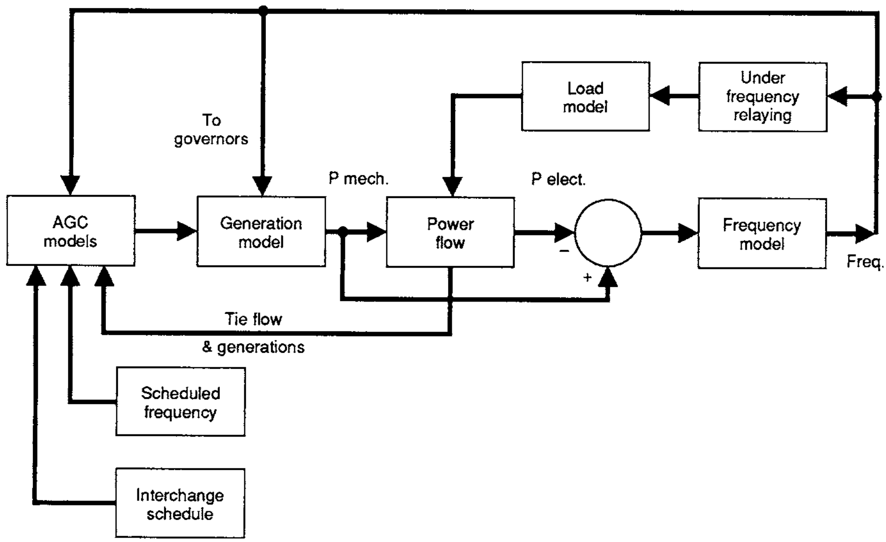


FIGURE 16-3 OTS control-response model.

runs last only a few seconds. The time frame for response of human control actions is the determining factor in the design of the simulation. Events that are beyond the range of human perception are not of interest, especially when viewed by telemetry with 10-s scans and through workstations with sampling of about 2 s. At the other extreme, it is important that the simulation be run in real time and be economical for runs of a half hour or more. These considerations result in an emphasis on prime mover dynamics and system frequency behavior in the structure of the simulation.

Because of the time response of AGC and operator control, we are dealing with low-speed phenomena rather than the transient and synchronizing effects not observed by the controller (either AGC or human). Also, because of the requirement for real-time response of the simulated power system, extensively detailed models of components with small time constants would require a short integration time step and a correspondingly heavy computational burden, so in this case we require a rather coarse time step (1 s) as compared with transient stability.

During steady-state operation conditions, line flows and losses are the result of generation, excitation, and load. The network solution is, therefore, more than adequately modeled by an efficiently coded load flow. A schematic of the control-response model is shown in Fig. 16-3.

16.2 RELAYING AND PROTECTION

By GUSTAVO BRUNELLO

The fundamental concept of protective relaying is to detect and isolate faults and other destructive phenomena in the shortest possible time consistent with economics and security. The principles vary at different points in the power system because of differing constraints. Distribution-system relaying must coordinate with fuses and reclosers for faults while ignoring “cold-load pickup,” capacitor bank switching, and transformer energization.

Transmission line relaying, on the other hand, must be sufficiently discriminating to locate and isolate any type of fault and do so with sufficient speed to preserve stability, to reduce fault damage, and to minimize the impact on the power system. This dictates the use of one or more pilot relaying systems.

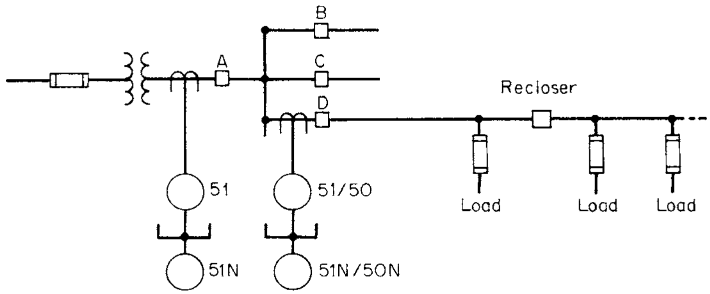


FIGURE 16-4 Typical distribution circuit relaying.

Subtransmission relaying varies from complete pilot relaying to simple directional overcurrent relaying depending on the importance and general nature of the subtransmission system.

Distribution-System Relaying. Typical distribution circuit relaying is shown in Fig. 16-4. Only one set of feeder relays is shown. This arrangement would be repeated for each feeder. The time-delayed phase and ground relays 51 and 51 N usually have a high degree of inverseness in their current-time characteristic to coordinate with the fuses and reclosers that are farther out on the circuit. The instantaneous units 50 and 50 N are typically set to trip the feeder breaker and protect the fuses when a temporary fault occurs beyond the fuse. For this type of fault, the feeder is removed from service by a reclosing relay that allows the fuse to blow when reclosing into a permanent fault.

The 51 N relay must be set with care to avoid its operation on loss of single-phase lateral load when a fuse blows. The "normal" load unbalance can be controlled to a reasonable degree by carefully supervising the balance of load connected to each individual phase (usually a 4-wire circuit with line-to-neutral connected loads). The opening of a fuse to clear a fault, and thereby drop load associated with one phase, will produce a much higher than normal load unbalance. This must not be allowed to cause operation of the ground relay. Its sensitivity is largely regulated by this consideration.

Cold-load pickup is the phenomenon whereby a feeder being reenergized after a long outage will experience a load appreciably in excess of maximum steady-state load (as a result of loss of diversity by thermostatically controlled devices). The feeder relays must ignore this if sectionalized reenergization is to be avoided. The relays on breaker *A* in Fig. 16-4 provide primary protection for the bus and backup protection for the feeder relays and breakers. In general, they are time-delayed and coordinate with the feeder relays with the accepted sacrifice of clearing speed for bus faults. These phase relays provide some measure of thermal protection for the supply transformer.

Modern microprocessor-based systems contain not only the instantaneous and time-delay relaying described above but, in addition, may contain reclosing, instrumentation, and fault data storage facility.

Subtransmission Relaying. Loops and multiple power sources used in feeding loads from the subtransmission system usually dictate the use of directional overcurrent relaying, distance relaying, or pilot relaying. In general, a subtransmission system is not intended to transmit bulk power from one location to another. Multiple sources are used purely in the interests of continuity of service.

Figure 16-5 shows an example requiring directional overcurrent relaying. A fault on the upper line would cause equal currents to flow in relays *A* and *B*. For this fault case, it is desired that relay *A* trip and *B* restrain. A fault on the lower line also causes equal current to flow in relays *A* and *B*. For this case, it is desired that relay *B* operate and relay *A* restrain. These two cases produce requirements that are mutually exclusive using simple overcurrent relays. The requirements can be met with directional overcurrent relays. If directional, the *A* relays would respond only to faults on the upper line and the *B* relays only to faults on the lower line. Coordination between *A* and *B* then becomes unnecessary.

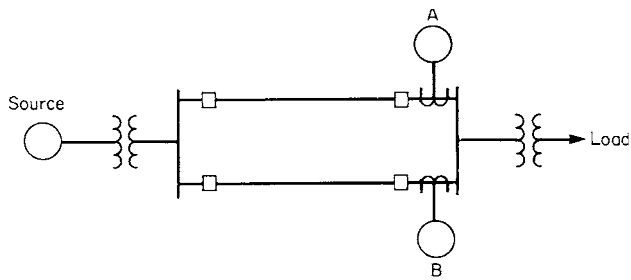


FIGURE 16-5 Partial one-line diagram of typical subtransmission system showing locations where directional relays are required.

Figure 16-6 defines in the simplest form a criterion for establishing where directional overcurrent relays are desirable.

Relay *R* in Fig. 16-6 requires consideration of distinctly different criteria, depending on whether instantaneous or intentional time delay tripping is involved. An instantaneous device at *R* must be set in such a way that it will never respond to a fault beyond bus *B*. The setting will be dictated by the maximum fault contribution (phase-fault contribution for phase relays or ground-fault contribution for ground relays or *phase relays*) for a fault at *B* and by the influence on the measuring unit of the dc component in the fault current. For example, a maximum fault at *B*, producing 20 A in relay *R*, would require a setting in excess of 20 A. If the maximum overreach factor for the particular instantaneous unit in use were 1.3 and a 10% margin were desired, a setting of 1.3 (1.1) (20) = 28.6 A would be required.

If a reverse fault such as a fault near bus *A* on other circuits could cause current in relay *R* to exceed 20 A (symmetrical), a higher setting would be required for this instantaneous unit than 28.6 A because the same overreach and margin factors would apply.

Since the extent of line coverage is dependent on the setting of the device as well as the source-line impedance ratio, a reverse fault which dictated a higher setting would cause the extent of line coverage to be smaller. By using directional control, no consideration need be given to reverse faults.

If the magnitude of relay current for this maximum magnitude reverse fault were less than 20 A, no consideration need be given to the inclusion of directional control for the instantaneous unit. A nondirectional relay will be satisfactory in this application because the relative fault currents make the relay inherently directional.

Time-delay overcurrent relays differ in their criteria from those of the instantaneous unit. In the interests of backup protection, relay *R* should always be able to detect the minimum fault on and beyond bus *B*. Further, in any time-delay relay applications, this minimum case should produce an adequate multiple of pickup current in the relay to ensure a clearly predictable operating time.

If, for example, the minimum fault at *B* produced 14 A in relay *R*, a setting of 7 A would be required (to give a multiple of pickup of 2 for this minimum fault case). If a reverse fault could deliver current sufficiently large to cause operation of a relay set at this level, consideration should be given to the use of directional control of the time unit. A frequently used conservative summary of this concept is that if the maximum reverse fault current can exceed 25% of the minimum fault current at the next bus, use directional control.

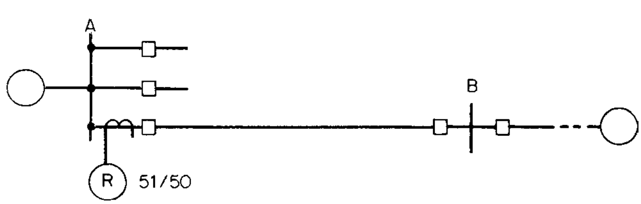


FIGURE 16-6 Directional relaying criterion.

The combined criterion for these concepts is—use directional control if a reverse fault could influence the sensitivity of relaying used to detect forward faults or if selectivity would not otherwise be possible. If source variations restrict instantaneous coverage to less than 50% of the protected line, or if the tripping times realizable for time-delay relays become undesirably long, distance relays should be used.

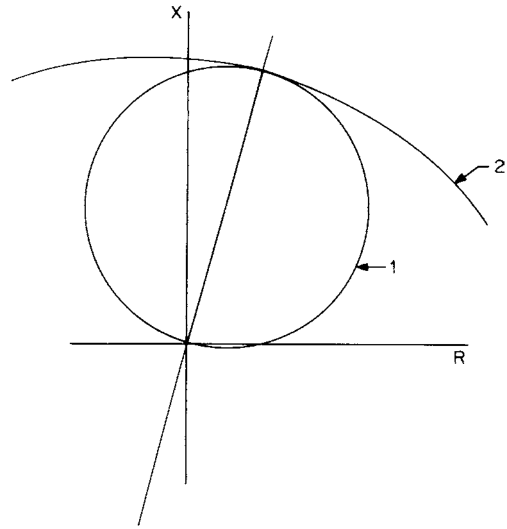


FIGURE 16-7 Resistance-reactance plot of distance relay characteristics.

Distance relays respond to the voltage and current applied to them and are usually more highly responsive at some lagging current angle. Figure 16-7 shows a typical *R-X* diagram that describes the behavior of these devices. Most distance relays in current use, phase and ground, have a characteristic similar to curve 1 or curve 2. Faults producing an apparent impedance at the relay location that falls inside the characteristic circle will cause the relay to operate. Since a distance relay has a distinct "reach" irrespective of source impedance and is directional, it is said to protect a "zone." Zone 1 relays are set to cover a portion such as 80% to 90% of a subtransmission or transmission line. Zone 2 relays respond to faults at all locations on the line and also to others in proximity of the line end. This is shown in Fig. 16-8. Zone 2 relays are typically set to cover 100% of the protected line plus 25% to 75% of the shortest line departing from the remote bus. Since they overreach the next bus at the end of the protected line, they must have a time delay or be associated with a pilot relaying system in order to preserve selectivity with other relays. A zone 2 relay should not be set to overreach any zone 1 relay at the next forward station.

A zone 3 relay is also often used and may be directional in the same sense or opposite sense as the zone 1 and zone 2 relays, or in some applications may be nondirectional. Figure 16-8 shows a one-line diagram with a "reverse-looking" zone 3 relay. The user shall carefully verify that the impedence reach for zone 3 is less than the load impedence presented to the relay under the most unfavorable steady-state operating conditions (overhead and overvoltage) of the system.

Microprocessor-based distance relay systems provide multiple zones, complete phase are ground distance protection, plus pilot logic, instrumentation, fault-data storage, and oscillographic information. However, in the past, simplified distance-relaying schemes were sometimes used in the interests of economy. One type used a complete complement of relays for one zone, which was initially set for a zone 1 function. A "starting" unit (overcurrent or distance) used to sense the presence of a fault. After a time delay, the setting (reach) of the relay was extended to zone 2 and still later to zone 3 (forward). A further abbreviation of this scheme allowed the starting units to identify the type of fault and to connect the appropriate voltages and currents to a single distance unit. These systems vary substantially in complexity, redundancy, dependability, and cost. The choice of one system over the

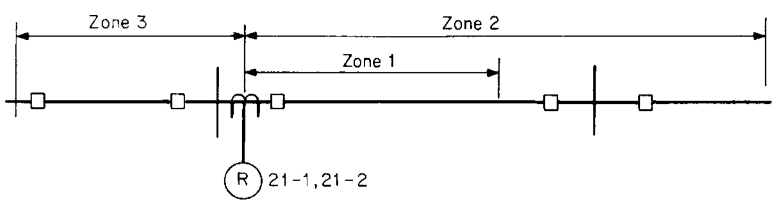


FIGURE 16-8 One-line diagram showing concept of distance relay zones.

others is dictated by the relative importance that is placed on each of these factors and the significance of the compromises involved in making such a choice.

Transmission Line Relaying. High-speed clearing of faults is universally required on transmission systems in the interests of maintaining stability, minimizing disturbance to wide areas of the power system, and decreasing fault damage. Pilot relaying is an important ingredient in this process. Pilot relaying entails the use of information obtained from one or more remote terminals as well as local information to establish the need to trip (or refrain from tripping) a local breaker. The remote information is transmitted by power line carrier, microwave, tones, pilot wires, optical fiber, or some combination thereof. An abundance of pilot-relaying systems are in use, each having its individual strengths and marginal weaknesses and each having varying degrees of dependence on the integrity of the channel.

Pilot Channels. Figure 16-9 shows one of the many types of pilot channels in use. This particular arrangement uses "power line carrier." The pilot channel is chosen sufficiently higher than the power frequency to allow separation to be achieved easily, generally 30 to 300 kHz.

Types of Protective Relaying Systems. Two basic systems form the nucleus for the families of pilot-relaying systems applied to transmission lines. They are the directional-comparison and the phase-comparison systems.

Directional-Comparison Relaying. The fundamental concept of the directional-comparison system is shown in Fig. 16-9. A directional relay at *A* responds to faults to its right as shown by the directional arrow in the figure. A similar relay at *B* responds to faults to the left of *B*. Both relays respond simultaneously only to faults on the protected line. The communication channel informs *A* about the state of *B*, and another informs *B* about the state of *A*.

One-to-one and a-half-cycle initiation of tripping is commonly achieved at both terminals following the occurrence of a fault on such a protected line. No tripping of these relays occurs for faults on other line sections. Abbreviated descriptions of the commonly used directional comparison schemes follow.

Directional-Comparison Blocking. In this system, each of the terminals is equipped with tripping and carrier-starting relays. The tripping relays are directional toward the protected line and are set to respond to all faults on the protected line and 25% to 50% beyond. This is called an overreaching setting. The carrier signal is required to prevent tripping for faults in that 25% to 50% overreaching area. Tripping at *A* is blocked by a signal transmitted from *B* and received at *A*. Transmission of the signal is initiated by a carrier-starting relay that operates for faults outside the protected line section.

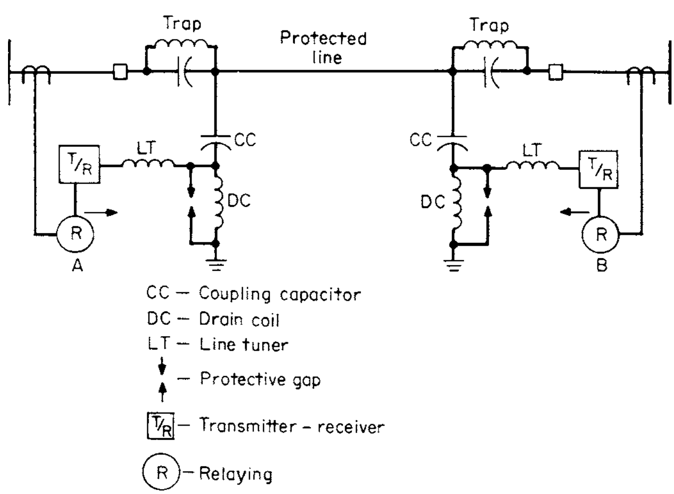


FIGURE 16-9 Representative channel for pilot relaying.

Internal faults are cleared by the tripping relays at all terminals, which have overriding control to stop all carrier transmission. A single-frequency on-off carrier may be used for both directions of transmission (A to B and B to A) because all carriers are turned off for an internal fault.

Underreach Blocking. This system uses a zone-extension scheme to limit, in the interests of economy, the number of distance units required. A relay set to cover zone 1 (the area from the relay location out to 80% or 90% of the protected transmission-line length) is stepped, after a coordinating delay such as 4 ms to zone 2 reach (covers the entire line) provided blocking carrier is not received from other terminals. If carrier is received, zone extension is still carried out, but at a much later time (often 15 cycles), to provide backup coverage for remote bus line sections and apparatus. Different carrier frequencies are required for the two carrier channels. Station A carrier cannot be allowed to block station A tripping because carrier cannot be stopped for some internal faults.

Acceleration. Zone extension is again used with this system. A frequency-shift carrier channel is preferred because transmission through a fault on the protected line may be required. A guard frequency is transmitted during nonfault conditions. The protective relays are given a zone 1 setting. All faults on the protected line are seen by one or both of the relays at the two ends of the line. Each causes carrier to be shifted to a trip frequency.

Receiving trip frequency causes the zone 1 setting of each local relay to be extended to zone 2 distance immediately. All faults in the area of overlap of the two zone 1 settings will be cleared without regard to the carrier signal. End-zone faults (faults not covered by the zone 1 relays at one of the terminals) will be cleared at high speed and essentially simultaneously once zone 1 extends to zone 2 reach.

Permissive Transfer Trip. In a permissive scheme, tripping occurs when the distance relay operates at each terminal and a trip signal is received at that terminal. The distance relays at the two ends of the line cooperate to clearly identify a fault as being "internal" to the protected line or "external." Permissive transfer-trip relaying systems are identified as *overreaching* or *underreaching* system, depending on the setting of the directional distance relay that keys the frequency shift tone or carrier transmitter at each line terminal.

If the system has a setting that causes it to respond to faults on the protected line and additionally to faults beyond the end of the protected line, it overreaches the remote relay, and the system is identified as an *Permissive Overreaching-Transfer-Trip (POTT)* system.

Underreaching schemes have the distance relays set to respond to faults within 80% of the protected line length. When they operate, they key the frequency-shift channel transmitter from "guard" to "trip" as well as immediately tripping the local breaker(s) without regard to action at the remote terminal. The two categories of these systems are identified as *direct* and *permissive.*

In the *Direct-Underreaching-Transfer-Trip (DUTT)* system, receiving the channel trip causes tripping of the terminal breaker(s). No local fault-detector relay operation is required. Strictly speaking, the direct scheme is not a directional-comparison system, because operation of the zone 1 relay issues a command to trip all breakers associated with the protected line, and no comparison takes place.

In the permissive underreach scheme, a local directional distance element, that overreaches the remote terminal, is required to supervise the tripping. Each terminal has two measuring elements: a zone 1 distance that underreaches the remote terminal and a supervisory element that sees faults beyond it. This scheme is called *Permissive-Underreach Transfer Trip (PUTT).*

Note that permissive transfer-trip systems require that a signal be received by the channel equipment in order for tripping to take place. These systems are usually committed to channels that are not dependent on the integrity of the protected power line itself such as pilot wires and microwave.

Unblock System. The unblock pilot relaying system is virtually identical to the overreaching-transfer-trip system but contains provision for allowing short time (100 to 150 ms usually) tripping when the channel fails, provided a local overreaching distance relay operates. Trapping of the transmission line prevents "loss of channel" from occurring on external faults. Loss of channel not accompanied by operation of a distance relay merely sounds an alarm to indicate that condition.

Each of these schemes represent varying layers of complexity imposed on the basic concept of allowing one or more distance relays at each terminal to identify the existence of and the direction to a fault. Use of the pilot channel allows the two terminals to share this information and to initiate the appropriate action based on the comparison. While the description is in terms of 2-terminal applications, they may in general be applied to the protection of 3-terminal lines. These systems incorporate subtle differences and small variations in their levels of security and dependability. They do differ in cost and capability, and their choice is greatly influenced by personal choice and individual previous experience.

Phase-Comparison Relaying. This form of pilot relaying compares, over a communication channel, the instantaneous direction of current at the two ends of the transmission line. To allow the use of a single channel, some such systems use a combination of the individual phase currents to generate a single-phase quantity for comparison. Others use a combination of the symmetrical components (positive, negative, and zero sequence) of the phase currents, and by applying appropriate weighting factors to each and adding the combination, a single-phase sinusoidal voltage is produced and converted to a square wave for comparison at the two terminals.

The concept of the scheme is that external faults will cause the local and received remote quantities to be essentially equal in magnitude but opposite in direction, while internal faults will cause them to be possibly different in magnitude but essentially in phase. In the comparison, the local quantity is delayed by an amount equal to the inherent channel delay, providing near-perfect coincidence for external faults.

The *segregated-phase-comparison* system compares the instantaneous direction of current at the two ends of the transmission line for each phase rather than utilizing some weighted combination of the currents or their symmetrical components. Modern high-speed channels allow information related to four subsystems (3 phases and ground) to be transmitted over a single voiceband in each direction. A local sinusoidal voltage proportional to phase current is converted, for each phase, to a square wave delayed by an amount dependent on channel delay and compared to the received remote quantity for the corresponding phase. Internal faults will produce essentially in-phase comparisons. External faults will produce comparisons essentially 180° out of phase. Considerable angular variation in these comparisons will still provide precise information regarding fault location. The ground comparison uses $3I_o$ current at the two ends of the transmission line.

Current Differential Relaying. To acquire the advantages of differential relaying for transmission lines similar to those obtained for generators and transformers, a scheme is in use that allows the waveform at each transmission-line terminal to be made available at the other. By using pilot wires, fiber optic, a microwave, or multiplexed digital channels the information is transmitted to the other terminal from which a phasor quantity is derived for comparison to the local quantity (delayed by the appropriate amount commensurate with channel time). This is accomplished using all the various technological forms: electromechanical, solid-state, and microprocessor. Excellent sensitivity and speed ($1^1/_2$ cycles) are achieved with this system and because of the abundant availability of digital communication channels, current differential applied to transmission lines becoming very popular.

Generator Relaying. Generators are a vital part of a power system, and their protection deserves is critical consideration. For the larger machines, 50,000 kW and above, a consistent pattern of protection has evolved. For the smaller machines, economics usually dictates that greater risks be accepted.

Large-Machine Protection

Hazards. The hazards against which protective devices guard are faults, unbalanced currents, loss of field, field ground, instability, and other miscellaneous phenomena that will be described later.

Phase Faults. Phase-fault protection is invariably provided by differential relays as shown in Fig. 16-10. By using identical ratio and accuracy-class current transformers, any "through" phenomenon such as load, external faults, or power swings will produce essentially equal restraint currents I_{R1} and I_{R2}. For external faults, operating current I_{OP} will be the difference of the two ct (current-transformer) error currents, or zero in the case of equal or negligible errors.

Internal faults generally will cause I_{R2} to reverse with respect to I_{R1} and I_{OP} to equal the transformed total fault current. The relays that are usually applied here have a sensitivity that is dependent on the restraint. For high through current, restraint is high, and the required I_{OP} is high, thereby restraining properly for possible high differences in error currents. For low internal fault current, restraint is much lower, and the I_{OP} required is much lower, allowing sensitive detection of the fault.

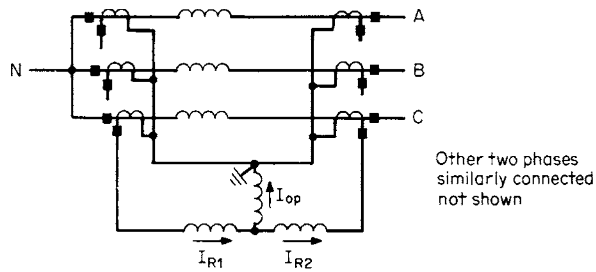


FIGURE 16-10 Typical differential protection for generator.

With this concept, large differential currents during external faults are ignored and the relay is sensitive for small differential current of internal faults.

Ground Faults. Stator faults involving conductor contact with grounded elements may cause essentially no current flow or current comparable to phase-fault levels, depending on the system neutral grounding. Most large machines are unit-connected, meaning the turbine, the generator, and the transformer are treated as a unit, with no fault switching at generator voltage level. The low-voltage winding of the unit transformer is delta-connected, providing zero-sequence isolation from all other segments of the power system. The generator neutral is grounded through a high-impedance circuit, usually a distribution transformer loaded with a secondary resistor. This combination limits ground-fault current to a few amperes, which is undetectable by the generator differential relay. With this widely used grounding method, the generator neutral shift is dependent on fault location. A ground fault at a generator terminal will cause full line-to-neutral voltage to exist between neutral and ground. The closer the fault to the neutral, the lower is the magnitude of this voltage.

A relay connected across the secondary terminals of the distribution transformer will be able to detect this voltage. It can be given sufficient sensitivity to detect faults from the line terminal down to approximately 4% of the neutral. It must ignore the normal third harmonic voltage, neutral to ground, to achieve this sensitivity.

The protection just described is blind to faults very close to the neutral point and consideration shall be given to complement with other relays or replace it with another principle. These schemes use the third harmonic voltage neutral to ground and sense its absence for a neutral-to-ground fault, or they interject a current at another frequency and supervise its level. Neutral-to-ground faults rarely occur and, in themselves, are of no consequence. A second ground fault not only will go undetected with neutral-to-ground fundamental-voltage-detection but also may destroy the generator.

Unbalanced Faults. Inherent in unbalanced faults is the fact that negative-sequence current is present. Flux associated with negative sequence rotates in a direction opposite to rotor rotation. This causes appreciable current flow in rotor structural parts that are not designed for such current, and excessive heating occurs. A relay designed to respond in a similar way to the machine is applied for this protective function. It is $I_2^2 t$ responsive, where I_2 is per-unit negative-sequence current (on the machine full-load current base) and t is time in seconds. Generators vary in capability from $I_2^2 t$ of 5 to 40 for negative-sequence currents in excess of full load, depending on the type and size of machine.

The negative-sequence current relay protects the generator against a prolonged contribution to an unbalanced fault beyond the generator breaker. It often contains provision for "alarming" at a lower level than the tripping level to annunciate the hazard of a sustained unbalanced current condition.

Loss of Field. Field failure caused by any event, such as loss of regulator, opening of field breaker, field short, or field open, will cause a large var flow into the machine and generally a substantial reduction in terminal voltage. This may or may not seriously jeopardize the machine, or it may jeopardize the stability of other adjacent machines. It requires detection and removal of the machine from the system.

Most loss-of-field devices utilize generator terminal voltage and phase current to obtain impedance and phase angle. Loss of field causes impedance at the relay to decrease and current to lead more. This

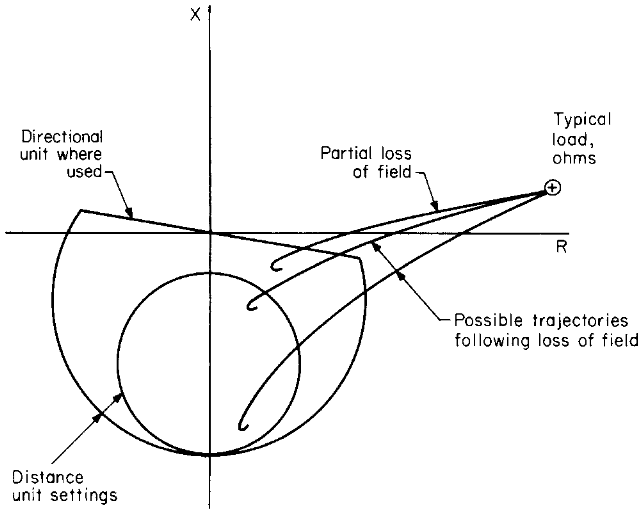


FIGURE 16-11 Detection of generator loss of field by measurement of impedance.

phenomenon is usually detected by "distance" relays as shown in Fig. 16-11. Apparent ohms as viewed from the machine terminals enter the characteristic circle of the relay, causing it to operate.

All such relays are equipped with time delay to avoid undesired tripping on power swings. Some contain directional and undervoltage units to permit additional sensitivity to partial loss of field and allow coordination with regulator minimum excitation units, the machine capability curve, and the steady-state stability curve.

Field Ground. A single field ground causes no machine distress. Allowed to go uncorrected until a second field ground occurs, it can cause sufficient magnetic unbalance to produce *catastrophic* vibration. For "brush-type" machines, detection of the first ground is usually accomplished by detecting current flow in a high-impedance dc-measuring circuit to ground. AC is also used in other devices through the introduction of an ac voltage between the dc field circuit and ground and monitoring the low-magnitude normal current that is allowed to flow.

Where a "brushless" arrangement is used, no normal access exists to the field circuit because there are no nonrotating parts at field voltage level as there are in brush-type machines. Monitoring for grounds is achieved by periodically dropping, manually or automatically, pilot brushes onto collector rings provided for the purpose. One collector is connected to the neutral of the 3-phase ac exciter, and the other is connected to the rotor structure itself. Measurement of the voltage between these two points with an overvoltage relay allows detection of a ground fault at any point in the field circuit.

Instability. When the electrical center appears to be in the transmission system, distance relays applied to protect the transmission lines can be used to detect instability and to separate the two system parts. This usually can be done discriminatingly with out-of-step blocking at some locations and tripping at others, all done in the interests of maintaining as nearly as possible a generation-load match after the separation.

On the other hand, when the electrical center falls in the unit transformer or in the machine, the normal complement of relays applied to generator or transformer protection either will not detect the out-of-step condition or will be time delayed to the point of being unreliable for this function. In these cases, out-of-step relaying is applied.

Figure 16-12 demonstrates the system behavior for a fault condition and for an out-of-step condition as viewed from the machine terminals and plotted in terms of a resistance-reactance diagram. Advantage is taken of the fact that emergence from the area between the blinder lines is on the same

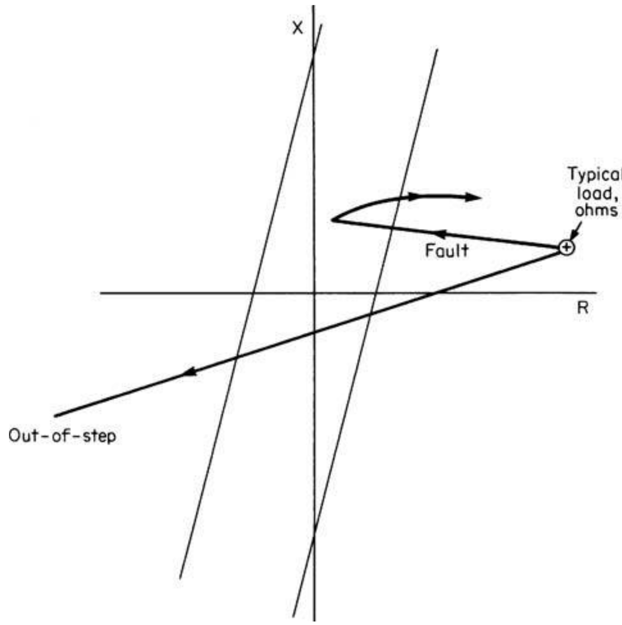


FIGURE 16-12 Blinder scheme for generator out-of-step detection.

side as entry for normal fault clearing and on the opposite side from entry for an out-of-step condition. A blinder-type out-of-step relay trips for the latter case.

Other Protection. For large, important units, relaying is included to detect motoring of the generator, inadvertent energization when the machine is at standstill, excessive volts per hertz that in turn causes excessive transformer and generator iron heating, stator and field overcurrent, and any malfunction not detected by the first line relaying (i.e., backup must be included to prevent catastrophic failure in the event of protective device malfunction).

Small-Machine Protection. Much individual preference goes into the choice of protective equipment for small machines. For the very small, only voltage-restrained or supervised overcurrent relays may be used. In some cases only over- or undervoltage and frequency detection is applied. In other cases, protection approaching that for larger machines is used.

In some cases, compromises with the more elaborate protection are used. For very small machines, time-delayed overcurrent relays with insensitive settings are used in the differential configuration. Specially connected watt relays are used for a combination loss-of-field and out-of-step detection function. Modern microprocessor packages contain most or all of the relaying functions necessary for generator protection plus monitoring, fault recording, and oscillography. They provide very low burden, self-checking, and greatly reduced panel-space requirements.

Motor Protection. Both synchronous and induction motors have protective requirements similar to those of generators. One important difference is that motors are accelerated by applying full or reduced voltage to their terminals, while generators are brought up to speed by their prime mover before being connected to the power system. Large starting current, then, is a normal expected phenomenon associated with motors that generators do not experience. Both types of devices contribute to external phase faults. Motor neutrals are not generally grounded, so no ground current will flow in an unfaulted motor.

Any protective device applied to protect a motor must ignore the conditions of starting current, load, and "through-fault" current, at the same time being able to sense low-magnitude internal-fault current. Differential relays perform this function well, often using a through-type current transformer with the two leads associated with each phase physically inserted through the ct window. Equal in-and-out currents generate no secondary voltage, so no operation of the relay connected to the ct secondary occurs. Internal faults cause unequal currents which generate a secondary voltage to cause instantaneous relay tripping. For larger motors, differential relaying schemes identical to those used for generators are used for phase-fault detection.

A ground-relaying variation of the "through-type ct" scheme requires that all 3-phase conductors be inserted through the ct window. Only ground faults on the motor side of the ct can cause the relay to operate. This is a widely used scheme. Another important element for detecting a fault in a motor is an instantaneous-trip phase device. It must, of course, be set above motor starting current, but available phase-fault current magnitude usually will greatly exceed the starting current magnitude, and very effective use can be made of this inexpensive and simple device.

Thermal Protection. Motors are usually equipped with devices that detect and relieve motor overloading. These are either devices that experience a heating effect comparable with that of the motor itself and act accordingly or are relays that detect the temperature of a resistance-temperature detector (RTD) (through a measurement of its resistance) embedded between conductors in the stator slot. As the motor temperature increases beyond the allowable level, the RTD resistance rises, and tripping of the controller takes place. Modern digital relays provide sophisticated models for the thermal behavior of the motor that operates when the thermal capability is violated.

Locked-Rotor Protection. Neither of the relays used for thermal protection will, in general, protect a motor with a locked rotor. A time overcurrent relay receiving one phase current will normally perform this locked-rotor protective function adequately. In some special large-motor applications where permissible locked-rotor time is less than the required starting time, distance relays have been used successfully to run timers to protect for the locked-rotor condition based on a measurement of a combination of motor impedance and phase angle.

Unbalance Protection. Any degree of voltage unbalance at the motor terminals will manifest itself in the form of increased heating in the motor, well beyond that which could be predicted from the increase in stator current. This can be sensed by a relay which measures voltage unbalance or negative-sequence voltage. Buses that supply a large number of motors are usually equipped with this kind of protection. Phase-current magnitude comparison also has been used very successfully on circuits supplying a single large motor.

Synchronous-Motor Protection. Because of the unique characteristics of synchronous motors, they are usually equipped with loss-of-field and out-of-step protection. This is often provided by a relay responsive to volt-amperes at an angle representative of the var flow into the motor on loss of field. It also will respond on loss of synchronism if the rate of pole slippage is compatible with the relay operating time or if the relay has a delayed resetting characteristic.

Transformer Relaying. Protection of large transformers generally consists of differential protection, gas space or oil rate-of-rise of pressure, or gas accumulation detection plus time overcurrent relays for backup.

Differential Relaying. The differential-relaying concept is applicable to transformer protection in a manner similar to that for generator protection, but distinct differences exist. While current transformers having essentially identical ratios and characteristics are obtainable in generator protection, no such identity is possible with the ct's used in transformer protection. Inherently, they *must* have different ratios and probably will have quite different characteristics. Also, inrush current on initial energization and following external fault removal is a very real phenomenon that must be accommodated by the transformer differential relay. These two circumstances, different ct's and inrush, makes the transformer differential relay different from the one described for the generator.

In addition to the fact that "through" conditions such as load or external faults produce different currents on the two sides of the transformer (to cause equal ampere turns in the windings), for a wye-delta or delta-wye transformer there is also a phase shift between the line currents on the two sides. Further, the standard ratios of ct's (such as 1200:5, 600:5, 100:5) used on the two sides of the transformer do not generally produce equal secondary currents for comparison by the differential relays for through conditions. As a result of these considerations:

1. Delta-side ct's are connected in wye.
2. Wye-side ct's are connected in delta.
3. Balance of input currents in the ratio of as much as 3:1 may be done inside the relay.
4. Inrush current is distinguished from internal-fault current in most transformer differential relays by using all harmonics, a combination of harmonics, or second harmonic only for inrush restraint.
5. Restraint is produced in proportion to the magnitude of the through current causing the relay to be sensitive at low current where ct error is likely to be low and to be insensitive at high current where ct error will be higher.

Microprocessor relays are able to perform these functions, previously assigned to electromechanical and solid-state relays. They allow all the current transformers to be connected in wye, irrespective of the protected transformer connection, through the use of an algorithm that supplies the appropriate phase shifting. This permits retention of phase designations for the monitoring and oscillographic display.

A widely used scheme for protecting a wye winding of a transformer against ground faults is shown in Fig. 16-13. The auxiliary transformer is carefully chosen with a ratio that will minimize the effect of ct error for external faults and force a restraint condition (currents not flowing into the winding polarity markers simultaneously) to exist. Internal faults produce a reversal in the operating current direction with respect to the polarizing (reference) current direction causing the relay to operate. Another common application uses a time overcurrent relay supplied by a neutral ct connected in a wye-winding ground connection. It must be time-coordinated with other ground relays on the power system connected to the wye winding. Where differential relays are used, the primary function of this neutral ground relay is to back up these other devices.

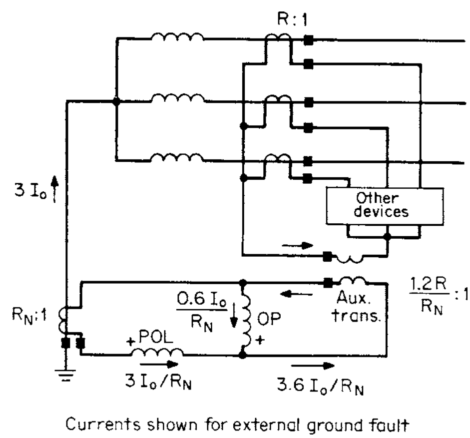


FIGURE 16-13 Transformer wye-winding differential protection.

A neutral-ground relay may accomplish a primary (or first-line) relaying function where low-resistance grounding is used and high-voltage fuses are used. The typical fuse size required for full-load capability will not detect a low-voltage winding failure to ground in such a case. The ground relay will, depending on fault-current level. Remote tripping of a breaker feeding the fused transformer will be required. Tripping of a low-voltage breaker will not clear this type of fault.

Rate of Rise of Pressure or Gas Accumulation. Depending on whether a transformer is designed to have a nitrogen space above oil or to have a "conservator tank" and be completely filled with oil, use will be made of a rate of rise of gas pressure or a rate of rise of oil pressure device in larger transformers. Normal load cycling causes pressure change, but the rate of change is moderate. Faults under oil cause a much higher rate of change, and this distinction allows this type of device to distinguish between load change and faults. Gas-accumulation relays collect any gas generated under oil by arcing or excessive temperature and base their fault detection on the extent of this collection.

16.3 POWER-SYSTEM COMMUNICATIONS

BY GEORGE R. STOLL

16.3.1 Introduction

Power-system communications play a vital role in the safe and efficient operation of the electric power grid. Real-time automation and control of electric utility generation, transmission and distribution systems are dependent upon reliable and secure communication networks. Through an ever-expanding role, the communications' networks enable the application of more computer and microprocessor-controlled devices. These networks and devices support the better utilization of extensive EMS and corporate information technology infrastructure. In addition, they enable the provision of new energy-related services and enhance the reliability and the safety of personnel and equipment.

Power-system communications application typically support various elements of a power utility's control, planning, accounting, and administrative functions. With deregulation of the U.S. electric utility environment, additional functions, including the marketing of bulk electric energy and transmission line access is also required.

This section will primarily address communication functions and the systems they support for power-system *operations*. The operation functions have certain communication requirements that are unique to the electric utility industry. Many of the other utility telecommunication functions, such as administrative voice and data, have communication requirements similar to those of other large business enterprises.

16.3.2 Communications/Control Hierarchy

Most power systems are vertically integrated; they perform the functions of generation, transmission and distribution, typically owned by the same entity. With deregulation, this model is in transition and ownership, and operations responsibilities vary with the country, the type of ownership and the size of the utility. In the United States today, the power marketing function (including the scheduling of generation, sale of bulk energy and transmission line capacity) has been separated from the transmission grid operations. This is to allow for an independent, nonbiased marketing of energy and transmission line capacity in a competitive market place. Previously, these functions were integrated into the same dispatch and operations center and frequently performed by the same personnel.

While deregulation is changing the way many companies are organized and requiring more communications infrastructure, the basic communications functions between the various control centers, generating facilities, transmission, and distribution elements are fundamentally similar. Figure 16-14 is a simplified overview of the current U.S. model, illustrating the relationship between major power-system elements along with their communication requirements.

16.3.3 Utility Communications Network Design Considerations

Most private utility optical and microwave wide area networks use time division multiplex (TDM) as the means for allocating a portion of a network's bandwidth to an individual circuit. By implementing a high sampling rate for each individual circuit, a very high quality, low delay (also termed *low latency*) channel can be transported over the private network. This high-quality and low-latency circuit format is well suited to the mission critical and oftentime sensitive circuit requirements of an electric utility. And, this is not unlike the common carrier and public network cable, wireless and optical transport schemes used until recently.

With the advent of the Internet and associated technologies, wide acceptance of a form of packet-switching protocol, termed Internet Protocol (IP), is starting to become widely used in wide area, public communication networks. Unlike TDM circuits, a packet circuit is not continuously connected. The packet circuit divides the information to be sent into packets, and each packet may take a different route through the network (or series of networks) to reach its destination. At the destination,

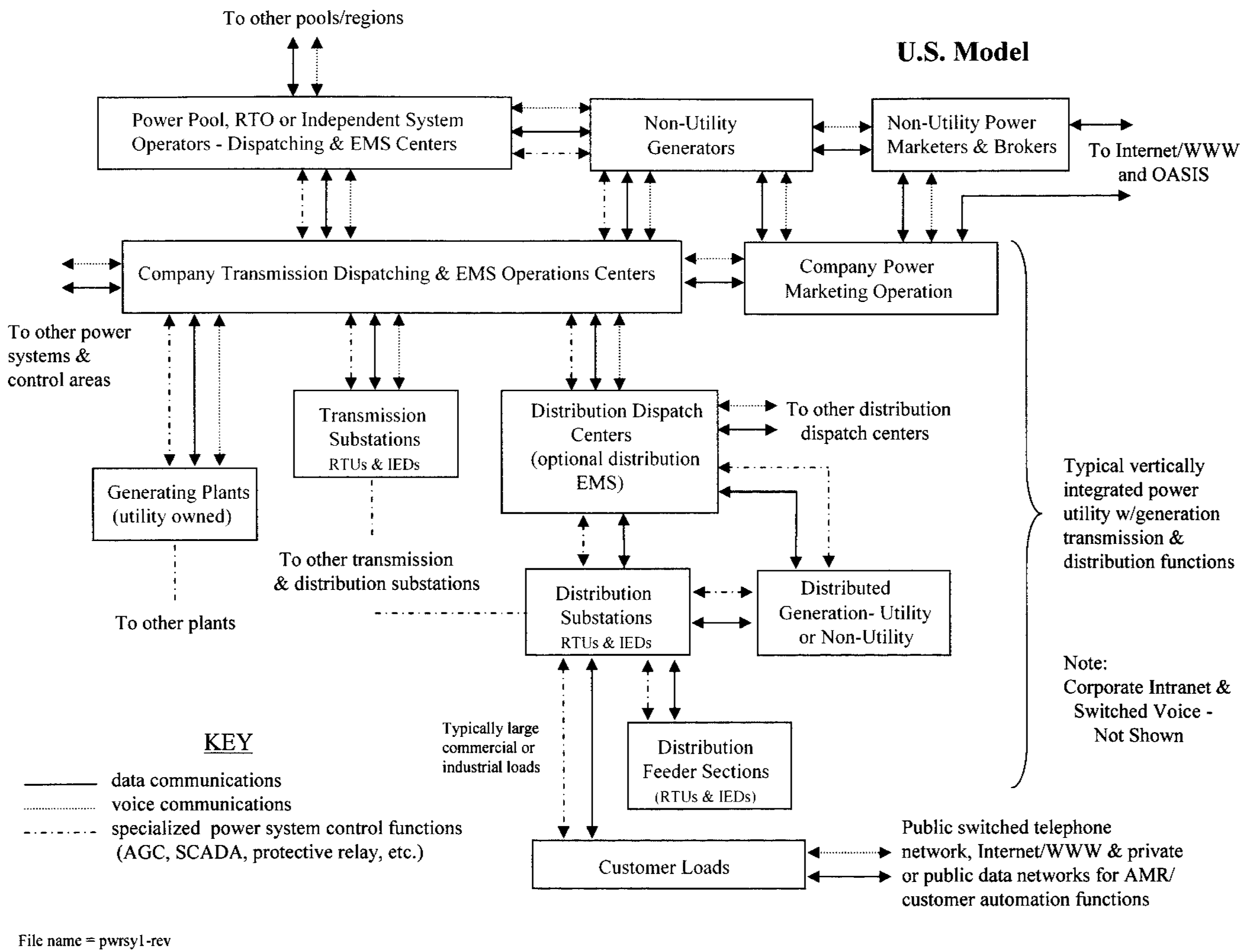


FIGURE 16-14 Interconnected power-systems—telecommunications requirements U. S. model.

the packets are assembled into a sequence matching the information originally transmitted. When no information is being transmitted, no network resources are utilized. This feature allows other users of the network to use its capability, or share the resource and thus increase its efficiency when used for less time critical data transmissions.

However, the disassembly, transmission time and reassembly of packets of information over IP networks add latency to all information being transmitted. Depending on the quality and capacity of the IP network and other technical factors, the information transmitted may be delayed anywhere from tens of milliseconds to several seconds. While acceptable for many types of data and even voice, these delays are not acceptable for many types of power-system operation channels.

This does not mean that IP cannot be used to implement operations circuits. With the use of significant additional bandwidth, IP formats can be transported with the latency and constant delay required. With excessive bandwidth, TDM circuits can be emulated over IP. However, this is typically only implemented over optical fiber networks where the utility manages the network. Because of the low cost of IP equipment and the availability of private utility company fiber networks with much more bandwidth than microwave, more IP networks for utility communication systems will be implemented in the future.

16.3.4 Specialized Power System Communications

In addition to the voice and data network communications typical in many multifacility industrial or business complexes, there are communication requirements unique to the electric power industry. These include:

- Protective relay
  Transmission line protection
  High-voltage switching equipment protection
  Generator and transformer protection
- Telemetering and telecontrol
  Analog and digital telemetering
  SCADA
  Remote alarms
  AGC
  Remote metering—real-time metering—revenue metering
- Voice communications
  Dedicated dispatch phones
  Two-way radio dispatch
- Other data communication links
  Computer to computer links
  Open access same time information system (OASIS) and the Internet
  Regional transmission organizations

16.3.5 Protective Relay Communication Channel Requirements

Protective relaying is unique in that the communications channel is faced with very stringent security and reliability requirements. Security requires that the communication channel never cause a false trip output. Reliability requires that the channel always be in service and function when needed. This requirement must be met even as the communication equipment is subjected to the harsh electronic environment present at substations, switchyards, and generating plants. Protective relay equipment must operate during and after a fault condition and in the presence of electrical noise, ground potential rise, and transient voltages common to these environments.

In addition, the communication channel must not add excessive time delay to the overall protective relaying function. Where the electrical circuit breaker is located at a remote location from the sensing relays, the channel is usually allocated up to 16 ms (about the time required for one cycle of the power-system frequency) total transit time. Extensive industry experience led the U.S. North American Electric Reliability Council to issue typical protective relay communication channel timing and redundancy requirements. Most protection schemes require communication channel times less than 16 ms. An exception to this is blocking schemes; requiring channel times less than 4 ms.[1]

Typical analog communication channel transit times range from 10 to 16 ms. These speeds are attained with direct links between the substations and usually via microwave or fiber communication systems. This is a function of the distance between the terminals, modulation methods, and the medium. An all-digital communication system can be faster than older analog transmission methods because there are no voice-frequency filters and baseband conversions, which add delays. However, in digital systems, application of digital access and cross connect functions or higher-order multiplexers can add unacceptable protective relay channel delays. Without the higher-order multiplexing and switching, one digital microwave equipment manufacturer reported a total channel transit time of 5.3 ms for a 640-km, all-digital microwave protective relay channel.[2] This channel transit time is inclusive of the transmitter time, the propagation time for the signal to travel to the remote end, the receiver detection time, and the operating time of the communications output device.

A majority of the total time required to clear a fault is a function of the breaker's operating time and the protective relay's detection time. These are those characteristics that cannot be readily changed. The communications channel is often the only variable element in the total timing required to detect and clear remote faults on the power system. With higher transmission line voltages, the longer the fault clearing time, the greater the potential damage to the utility's electrical equipment. Thus, the justification for the emphasis on protective relay communications channel speed.

16.3.6 Telemetering and Telecontrol

Telemeter and telecontrol signals provide information to the operators and may also serve as computer system input signals. Telemetering permits remote measurement of current, voltage, real and reactive power, position, flow, and other data relevant to operation of the power system.

Digital systems typically have an RS-232, RS-485, or an Ethernet interface to the communications media. Older analog systems typically use a transducer to convert the parameter being measured to a dc voltage or current. Telemetering equipment at the remote location linearly converts the dc voltage or current from the transducers to a sub-audible frequency, usually in the range of 10 to 30 Hz. This subaudible frequency is used to modulate frequency shift tone transmitters with an output in the range of 420 to 3300 Hz or higher. These signals can be transmitted over standard voice-grade communication channels carried on telephone, microwave, or optical fiber links. At the receiving end, tone receivers convert the audible frequencies back to voltages or currents (typically 0 to 100 mV or 0 to 20 mA), which in turn are used as inputs to monitoring, recording, control, or computing equipment.

With the emergence of sophisticated SCADA systems, a great deal of analog telemetry is being replaced with full digital systems or integrated into the SCADA system. The SCADA systems are designed to provide telemetry and control functions of multiple points (or subsystems) within a station. The SCADA's computer and electronics located at the remote location (such as a substation) are termed RTUs. The RTUs communicate with a master located at the dispatch or EMS center. Master and remote units communicate with each other using a series of digital messages that convey the addressing, control or status information and error checking. This communication can take place over a voice grade communication channel or via an all-digital communication link.

Supervisory control and data acquisition communication between RTUs and the master take place with one of three access sequences. These basic access methods include a polling format, a scheduled or a contention access.

Polling Access. The master periodically sends a request for information/control command sequence to the RTU.

Scheduled Access. The RTU initiates communications to the master on a predetermined schedule.

Contention Access. The RTU or the master may initiate a communications process whenever a control command needs to be transmitted or a status changes. The communications process includes the intelligence to perform data transmission collision avoidance, detection, and retransmission functions.

Combinations of these access methods is also common. In one variation, an RTU may not communicate with a master station unless a control status or data value has changed, termed *report by exception*. In other variations, certain critical events at the substation may trigger a normally polled RTU's immediate communication with its master station. Rather than wait for its assigned time slot in a polled sequence, it communicates virtually instantaneously with its master.

RTUs and Microprocessor Technology. The role of the RTU in the electric power substation is changing. The advent of intelligent electronic devices (IEDs) and programmable logic controllers (PLC) means more devices, such as protective relays, have electronic intelligence and the ability to be dynamically controlled and monitored. Rather than provide a dedicated communications channel to each IED or PLC, these devices may communicate with the local RTU. The RTU, in addition to its primary data collection and control function, acts as a data concentrator and protocol converter. It is becoming common place to see the IEDs and controlled devices networked together within a substation via a local area network (LAN).

16.3.7 Automatic Generation Control

Automatic generation control provides the telemetry and telecontrol to support tie-line and load-frequency control functions. These systems scan (sample) the individual unit generation and tie-line power flows. Via centralized computer control, they generate the raise/lower control pulses sent to individual electric generators. These systems can be very time-critical, but usually do not have as stringent a communications channel requirement as protective relaying. Formerly all analog and using dedicated channels, many of these systems today are now digital or incorporated into the EMS functions.

16.3.8 Voice Communications

Voice communication with and between field personnel and the various dispatch, power pool, and EMS centers takes place over telephone and radio systems. In addition to use of the public switched telephone networks, power systems frequently include dedicated circuits between the dispatch and control centers. Often these are configured so that minimal or no dialing is required. The circuit may be transported over private or dedicated networks. These are termed *hotline* or *ringdown circuits*. They allow dispatchers to communicate with each other without the normal 10- to 20-s delay caused by dialing and the public network's switching, routing, and signaling functions.

Two-way radio communications are used for communications with field personnel performing operations, maintenance, and electric service restoration. These systems operate in the very high frequency (VHF) 30 to 300 MHz or ultra high frequency (UHF) 300 to 1000 MHz portions of the radio spectrum. Larger utilities may use trunked radio systems, where multiple radio channels and their use is computer controlled. In a trunked radio system, the channels are dynamically assigned as needed, allowing efficient use of the radio spectrum. Smaller radio systems use conventional, dedicated radio channels. These are analogous to a "party line" environment, where all the users on the channel can hear each other. In these configurations, an individual radio channel may be shared by the various functions within the utility.

Large dispatch and energy control centers have radio dispatch consoles. These consoles consolidate all the radio system control functions, allowing the system operators to quickly access the mobile radio systems and to communicate with field personnel throughout their service territory. These radio consoles may be stand-alone units or have voice telephone functions integrated into them.

16.3.9 Other Data Communication Links

Bulk data information exchange takes place between the various power-system's computers. Data traffic over these links may include the EMS system's generation scheduling, fuel cost and generator availability, transmission capacity, load predictions, interchange billing, frequently weather information, and other data relevant to power-system operation. These data links are usually over dedicated communication channels and vary in data speeds from a single DS-0 (64 kbits/s) to T1 or E1 rates.

Oasis and the Internet. In the United States, the Federal Energy Regulatory Commission (FERC) established requirements for implementing open access to the electric transmission grid.[3] Ultimately, this open access to electric transmission lines will allow almost any generator of electric energy to sell to any purchaser. Implementation requires additional communications, establishment of wholesale power marketers and independent system operators who act independently of the utility that owns the transmission lines and/or generates the electric energy. To address how transmission line capacity and availability information would be made available to everyone at the same time, on a uniform basis, FERC defined an application of the Internet.[4] All entities that generate electricity for the open market, own electric transmission lines, or buy energy from these parties will need to use the Internet to exchange information. Termed OASIS, these Internet based, specialized information services allow anyone with access to the Internet to view the capacity, availability, and associated costs of electric power and its transmission. OASIS is currently serving as a shared database of information. It may develop into an intelligent system capable of performing generation scheduling, and control.

FERC Order 2000. Following implementation of the transmission line open access rules, FERC released an order requiring all the U.S. high-voltage transmission line owners to tell FERC how they would organize regional transmission organizations (RTOs)[5]. Regional transmission organizations are intended to be a consolidation of independent system operators (ISOs). The objective was to lower the number of ISOs and simplify the communications and marketing communications. Today, there are both RTOs and ISOs and the exact definition of the bulk transmission and generation model is still evolving. From a communications perspective, all variations of these power system models require additional voice and computer communication links along with more real-time data of power grid operations.

16.3.10 Communication Alternatives

As communication alternatives are considered for power-system operations, factors including circuit capacity, reliability, latency, jitter, and other technical parameters along with cost must be considered. Many organizations categorize their circuit requirements into *level of service* categories. Levels of power system telecommunication service can be grouped into three general categories.

System Critical. These are communication links that are extremely reliable and which support process and control functions requiring near real-time communications. Communication paths are available full time, they are usually dedicated to specific functions and the process or control function they support is usually computer controlled and has total response times ranging from milliseconds up to several seconds. Examples of system critical power-system communication links include protective relay channels, LFC, AGC, tie-line control, many SCADA systems, and some computer links.

System Priority. These are communication links that support voice and data functions with total response and control times ranging from several tens of seconds up to 1 h. Communication links may be dedicated or shared and provided by the power utility or a public carrier. The process or function supported may require or allow human intervention. The communication channels are usually very reliable and seldom blocked, or not available. Examples of system priority communication links include voice dispatch circuits, the voice two-way radio systems, computer-to-computer data links, and local and wide area data networks. They also include commercial and industrial electric load shedding, metering of large commercial and industrial loads, less time-sensitive SCADA functions, and distribution and feeder network automation systems.

System Administration and Support. These are communication links that support power-system functions where near real-time or very time-sensitive communications are not required. Channels support functions or processes with acceptable response and control exchanges ranging from minutes to days. If a communication exchange is interrupted or lost, it can be resent without severe implications to the safe and efficient operation of the electric power system. This category includes all types of administrative voice and data via the public-switched networks, private branch exchange systems and metering, planning, scheduling, billing, and customer service communications. In the U.S. model, it also includes use of the Internet for many of the transmission access scheduling and power-system marketing functions.

The level of service, process response times and media/provider selection criteria for typical power-system communications functions are represented in Fig. 16-15.

16.3.11 Communications Media/Service Type

Power-system communication networks are typically composed of systems using several technologies and, often, multiple service providers. Following are some of the most popular types of systems and services used to support the specialized needs of the power utility.

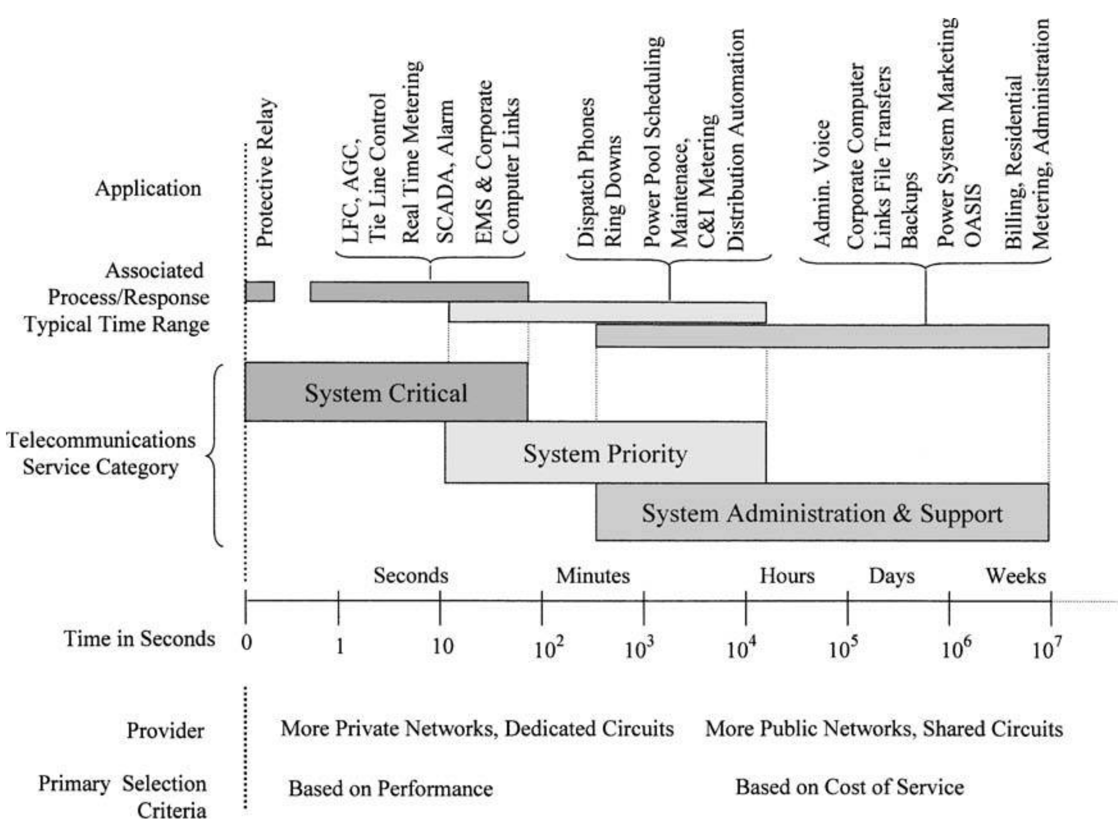


FIGURE 16-15 Power-system communications—application vs. service category.

16.3.12 Private Point-to-Point Microwave Systems

Private microwave systems have proven to be a very reliable and cost-effective method of supporting a wide range of communications needed for power-system operations. These systems operate in assigned frequencies ranging from 2 to 23 GHz. Very reliable links can be established for distances up to 80 km each, depending on the intervening terrain, the operating frequency and the height of the antenna systems. Multiple stations (repeaters) can be "chained" together for end-to-end system distances of thousands of kilometers. Most systems require licenses, although some spread spectrum systems with limited capacity and range are available for short-haul services.

A drawback to microwave systems is their requirement for a line-of-sight path between stations, often requiring large towers to support the antenna systems (Fig. 16-16). Spectrum may not be available

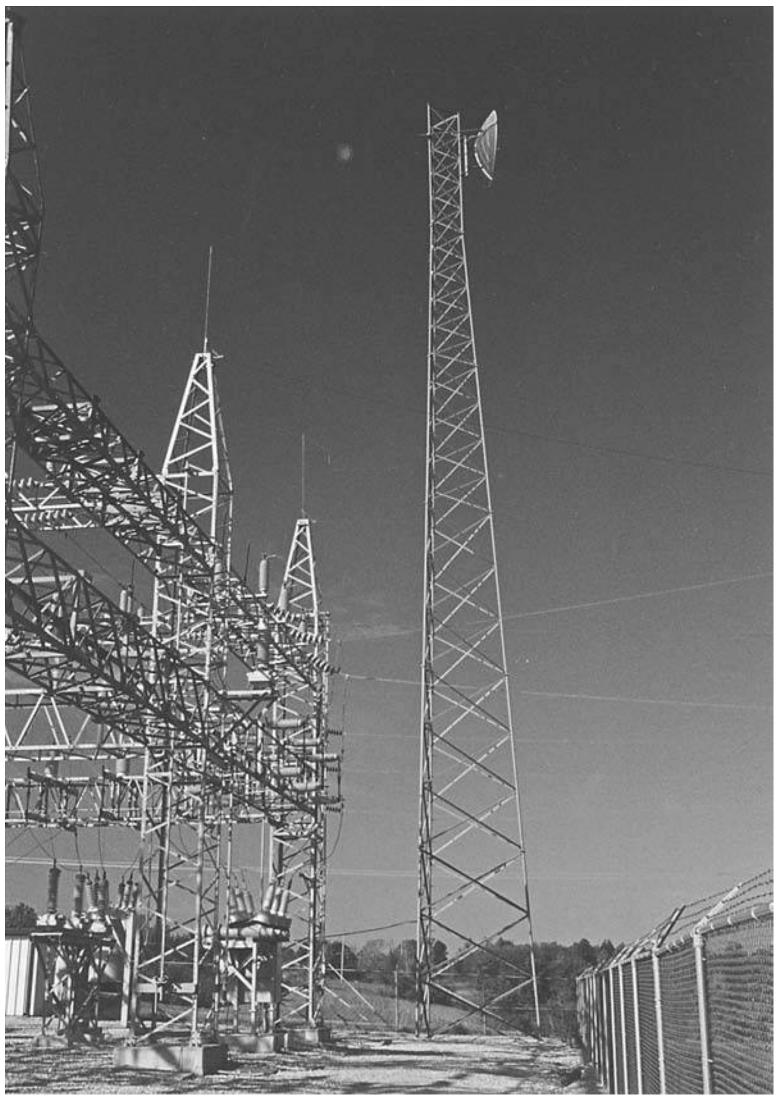


FIGURE 16-16 Microwave systems require large towers for line-of-sight paths between stations.

due to heavy use of the limited amount of assigned frequencies. Systems operating at 10 GHz and higher are more sensitive to rain and atmospheric conditions, resulting in shorter path lengths.

Digital point-to-point microwave systems can be designed to provide very high reliability with error rates of 10^{-9} or better. They also can be configured to provide minimum channel delay, so they are ideally suited for protective relay channel service. Once microwave systems are installed to provide required power-operations communications, it can be cost effective to use them also to transport the power system's administrative voice circuits and other corporate communications.

16.3.13 Leased Telephone Circuits

Telephone circuits, both dedicated, direct point-to-point, and the switched lines that are transported through the public telephone networks are widely used for power-system operations. They can be cost-effective for the system priority and system administration and support category functions that do not have extremely demanding availability, reliability, and error rate requirements. Switched circuits, also termed dial-up circuits, can be ordered and installed quickly in most populated areas. Error rates vary over a wide range and are a function of the quality of the local telephone company's cable plant and distances to central office equipment. Another concern with leased telephone circuits is that they are often transported by multiple carriers, making it difficult to identify intermittent problems.

Voice grade, dial-up circuits may be able to transport data rates to near 56 Kbps in metropolitan areas that are close to the telephone company's central offices. Rural environments can expect much lower data rates. Digital leased circuits are available in some areas that will transport data rates to DS-3 (45 mbit/s) or higher with good error rates, but these can be costly.

A majority of leased telephone circuits use a metallic conductor in some portion of the circuit, usually in the last mile segment where they enter and exit the power company facility. Their metallic conductor and cable sheath make them susceptible to induced noise and voltages (magnetic induction) and ground potential rise that are common in power-system environments. With ground potential rise, a local fault may cause the voltage potential of the electric station ground grid to rise to several thousands of volts, while the telephone company's central office ground remains at zero potential. This difference in ground voltage appears as a high-voltage on the utilities equipment connected to the telephone circuit. This cannot only damage equipment and the cable, but also presents a safety hazard to personnel working on the electronic equipment.

To protect the connected equipment and personnel in near proximity to these metallic cables, transmission substations typically require sophisticated telephone line protection devices. These are required to eliminate the harmful voltages that would otherwise be present on the telephone cables and the cable sheath. These devices may take the form of short fiber-optic links (with the fiber and all interface electronics mounted in a dielectric cabinet), isolating transformer or neutralizing transformer installations. The fiber-optic devices are generally replacing the neutralizing and isolating transformer installations because they do not require as careful a design or remote grounding considerations.

16.3.14 Satellite Services

Most traditional communication satellite systems use a single satellite placed in a geostationary orbit, 36,000 km above the earth's equator, functioning as a microwave signal repeater. Several newer systems use multiple satellites in lower orbits categorized as low-earth-orbit (LEO) and medium-earth-orbit (MEO) systems. Power-system operations have made only limited use of satellite technology.

Traditional satellite systems can transmit large amounts of digital information, in the form of voice or data, with low error rates. However, a characteristic of all geostationary satellites that eliminates them for most power-system-critical communications is the propagation time associated with the microwave signal. Transmitting the signal 36,000 km from the earth to the satellite and back again adds approximately 250 ms to the communication channel time. Many systems use a double-hop technology, where all signals are relayed through a large earth station, boosting the signal and thus allowing for the application of small parabolic antennas at remote stations (termed *very small* *aperture terminals,* or VSATs). This doubles the delay, yielding an overall channel propagation time of nearly 1/2 second. Add to this the access and back haul times, error correction and overhead, and round-trip times can be in the order of 2 to 3 seconds. This delay creates problems for the polling and response timing in many EMS systems and causes geostationary satellite technology to be totally unacceptable for protective relaying.

Other Satellite Services. Low-earth-orbit (under 2000 km from the Earth's surface) and medium-earth-orbit (10,000 km) systems are in service and are used in a limited manner for power operations that are not time critical (certain types of remote metering). These multiple satellite systems offer voice and data, usually on a worldwide basis. Low-earth-orbit systems are being used that can offer meter reading and some types of SCADA services. Because these systems are closer to the earth's surface, there is a considerable improvement in transmission delay times versus geostationary satellite systems. Primarily using packet data formats, there are still time delay issues.

16.3.15 Private and Commercial Land Mobile Radio Systems

Land mobile radio systems are widely used to support power-system field operations. They operate in the VHF region (30 to 300 MHz) or the lower portion of the UHF region (300 to 900 MHz). They primarily support voice operations over narrow bandwidth channels. Some systems also support limited mobile data and status messaging. However, with data rates over this bandwidth limited radio channel typically below 9600 bits/s, only a minimum amount of data and status information can reasonably be transmitted.

Many utilities have private systems (owned and operated by the utility) although a wide variety of commercial services are available. These systems use conventional (half or full duplex) communications over individual dedicated radio channels or for large systems, trunked access over multiple channels.

Commercial systems typically use more trunking and spectrum efficient systems and often cover wider service areas. They are seldom used for power-system dispatch operations because of concerns over reliability and channel access during busy periods or regional disasters.

Limitations of land mobile systems include the limited availability of additional spectrum, congestion on existing channels, and the requirement of an often complex licensing process. Radio propagation at UHF frequencies limits the system range to near line-of-site distances with somewhat greater distances for VHF. Installation of a wide area system requires locating the transmitting equipment on tall towers, buildings, or mountaintops.

16.3.16 Cellular and PCS Wireless Services

Cellular radio service (operating in the 800 to 900 MHz spectrum) and personal communications services (PCS) (operating in the 1.9 to 2 GHz spectrum) are used by power-system operations for the *system administration and support* category of mobile voice or low priority, low data rate (typically under 56 kbits/s) mobile data communications.

Cellular and PCS systems provide service in metropolitan and many rural areas. Low power mobile or handheld transceivers communicate with nearby radio base stations configured in a cellular pattern. The base stations (termed cell sites) are linked with each other and the public telephone network. The system is designed so that the base stations can reuse the radio channels of other nearby cells, thus allowing many subscribers to simultaneously use the spectrum.

Their circuits often do not meet the quality or reliability criteria needed for higher priority power-system operations. Serving as wireless telephones and frequently as low speed data transceivers, they are not designed with the same grade of service as the U.S. wire-line public telephone network. During busy periods, calls can be blocked from accessing the cellular network and calls in progress may experience interference or may be terminated.

In addition to the popular mobile voice communications, cellular service is also widely used in power-system operations to support low priority, occasional dial-up or unsolicited alarms from remote locations where more expensive, higher reliability EMS systems are not justified.

16.3.17 VHF and UHF Radio Data Links

Very high-frequency and UHF radio systems are frequently used to support EMS and telemetry links. These data-only systems can be configured as point-to-point circuits (typically in the VHF frequency region) or as point to multipoint (the UHF frequencies). Point-to-multipoint systems are also termed multiple address systems. Characteristics of most of these systems include data rates in the 2400 to 9600 bits/s range, but actual throughput is much lower than this, since most systems use a shared channel or some form of polling access. System error rates can be as high as 10^{-4}, but error correction and packet data transmission formats provide acceptable performance for many applications.

Limitations of these systems include lower data speeds and throughput rates and the requirement to obtain licenses for the radio channels. Most of the VHF and UHF radio spectrum allocated for this type of service is congested or assigned to mobile radio applications, making licensing in some areas difficult. They also require line-of-sight or near line-of-sight paths between the remote and the master station antennas, often requiring tall towers. Typical range of these systems varies from 10 to 40 km, which is a function of the operating frequency, intervening terrain, and height of the antenna systems. Despite their limitations, they are widely used in power-system data and SCADA systems because they can be designed to meet many *system priority* category communication requirements.

16.3.18 Power-Line Carrier

Power-line carrier (PLC) or carrier current systems transmit very low-frequency (65 to 300 kHz) radio signals over utility transmission and distribution wires. Although voice transmission is possible, most systems are used for telemetry and protective relaying. PLC systems provide a very reliable method of long-distance protective relay signaling for high-voltage transmission lines. More recently, distribution line carrier systems have been used successfully for automating distribution applications and automatic meter reading.

PLC systems can reliably transmit their low-frequency signals over transmission lines in excess of 200 km in length. Since existing transmission or distribution lines are used, no right-of-way or licensing is required. The primary limitation of these systems is their limited bandwidth, usually transporting two to four voice-equivalent channels. In addition, transformers and capacitor banks used for power factor correction severely attenuate the PLC signal.

16.3.19 Privately Owned Fiber Optic Cable Systems

Fiber-optic systems provide some of the highest quality transmission systems available with more capacity than any other telecommunications media today. Properly designed systems have extremely low bit error rates, on the order of 10^{-12} or better and capacities to 320 Gbits/s. Fabricated from very pure forms of glass, the hair-thin fiber strands are nonconductors, and thus not susceptible to the induced voltages and ground potential rise problems found in electric plant and substation environments. Fiber-optic cable is widely used for instrumentation, wide area networks and local area networks in and between the utility company's plants, substations, and offices.

Numerous utility companies are placing optical fiber cables (Fig. 16-17) in their electric-line rights-of-way. These can be used solely for internal communications, or excess capacity may be sold to other carriers. Specialized cable arrangements have been fabricated, suitable for the high-voltage transmission line environment.

There are several dozen variations of fiber-optic cable, protective sheath, and cable/messenger configurations used by utility companies on their electric transmission lines. A majority of the fiber being deployed on electric transmission line right-of-way today is in one of four arrangements. These include:

Direct Buried. An armored sheath protects a bundle of fibers. The sheath can be metallic or plastic.

Buried in Conduit. The conduit is usually nonconductive plastic or polyvinyl chloride style with two to four interior subducts. The multiduct conduit allows additional or replacement cables to be pulled into the duct system at a future date.

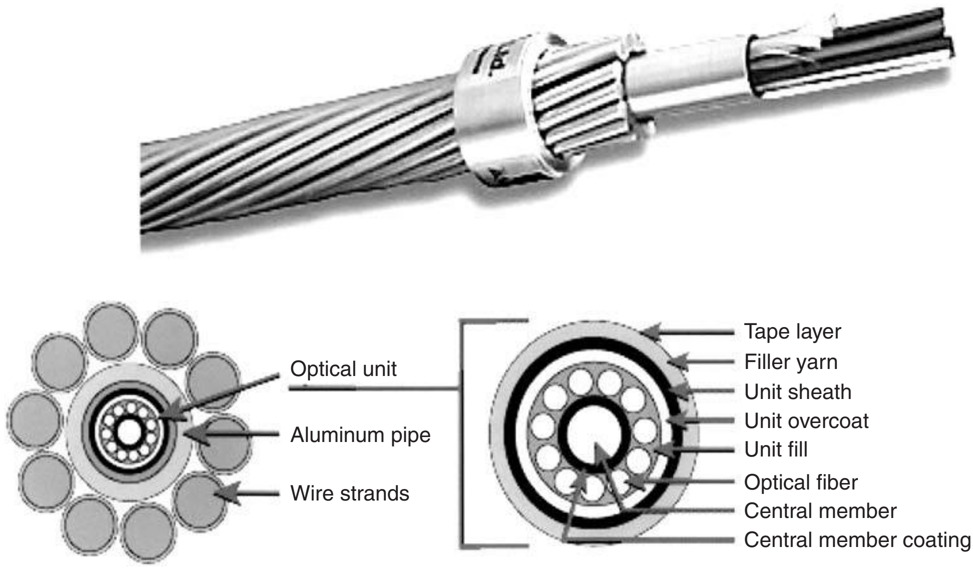


FIGURE 16-17 Optical ground wire.

All Dielectric Self-supporting Cable (ADSS). An aerial cable with a nonconducting protective jacket with Kevlar or fiberglass supporting members (Fig. 16-18). The nonconductive construction allows this type of cable to be placed close to the electrical conductors. This type of cable typically contains anywhere from 12 to 96 optical fibers.

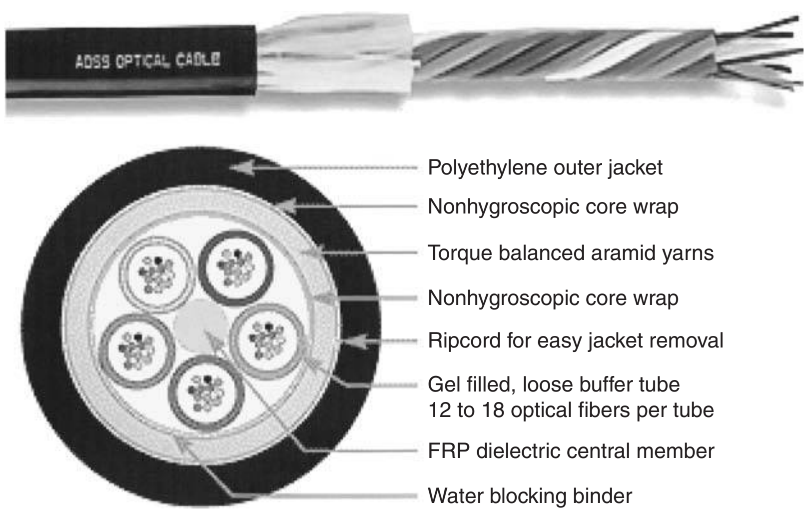


FIGURE 16-18 All dielectric self-supporting.

Optical Ground Wire (OPGW). The optical fibers are placed inside the metallic shield wire (also called a *static wire*) and suspended above the electrical conductors. In this application, the OPGW serves two functions: (1) it acts as a grounded shield wire protecting the electrical conductors from direct lightning strikes, and (2) it carriers the fiber-optic communication cables. This type of cable is available with up to 144 fibers. Fiber counts above 96 fibers are less common due to the physical weight of the added fibers and support materials.

These wide area network fiber-cable systems use single-mode fiber cable, allowing distances up to 100 km without repeater stations. New fiber-optic cables and the application of optical amplifiers allow even greater distances. Most systems in service today transmit one or two optical wavelengths in the 1310 or 1550-nm regions. New technology allows multiple optical wavelengths to be transmitted over an individual fiber (wavelength division multiplexing), significantly increasing existing and future system's transmission capacity.

REFERENCES

1 North American Electric Reliability Council: Planning Standards, Section III A., System Protection and Control, Section 3.2 Performance Tables, 1997.

2 Laine, R.U., and A. Ross Lunan: *Characteristics of Digital Microwave Links Supporting Utility Telecom Network Operations*, Technical Document No. 112, Harris Fairinon Division, May 7, 1993.

3 Federal Energy Regulatory Commission: Order 888, Open Access Final Rule, April 24, 1996.

4 Federal Energy Regulatory Commission: Order 889, Open Access Same Time Information Systems, April 24, 1996.

5 Federal Energy Regulatory Commission: Order 2000, Regional Transmission Organizations, December 20, 1999.

16.4 INTELLIGENT DISTRIBUTION AUTOMATION

BY DOUG STASZESKY

Supervisory control and data acquisition has long been used to control transmission systems to provide the operational flexibility and speed, required for efficient and reliable performance. The use of SCADA in the distribution system is becoming increasingly important as utilities move into a deregulated, competitive environment. The acronym SCADA has been generally replaced by the term distribution automation (DA), which incorporates the principle of operating switching, fault interrupting, and other control devices automatically in response to events in the system. Automated switching of distribution feeder circuits provides significant improvements in reliability, enhances operational flexibility, and increases the utilization of distribution assets and personnel.

Feeder switching and protection systems utilizing powerful IED's, sophisticated algorithms, a plethora of sensing devices, and all connected by increasingly fast and secure data communications enable the implementation of distributed intelligence, which is fundamental to implementation of an intelligent grid now and in the future. As DA supplanted SCADA as the term *du jour* for such systems, it is likely that a new term IDA—intelligent distribution automation—will come to represent the real needs of utility planners, engineers, and operators to meet long-term customer needs as well as the demands of regulatory bodies.

Just as mainframe systems are being replaced with flexible, fast-distributed computing networks made of PCs, centralized control of the distribution power system will move to distributed computing to become the intelligent grid. The intelligent grid will deliver benefits far beyond that, which can be delivered by conventional reclosers, switches, automatic sectionalizers, and other devices, which do not share information about the status of the grid. Distributing system intelligence effectively eliminates communication bottlenecks and time delays associated with more conventional, centrally controlled SCADA systems, and are sustainable even if single computing nodes do not function. And, when properly designed, systems based on distributed intelligence offer a completely scalable advanced feeder automation system that can easily, and cost effectively meet the challenge of the smallest tactical reliability problem, or grow to deliver system-wide automation functionality and improved asset utilization.

Distributed intelligence will become increasingly important as distributed energy resources (DERs) are deployed on the distribution system. Distributed energy resources, fully incorporated into the intelligent grid will further enhance reliability and power quality and have the potential to significantly improve overall asset utilization. Distributed energy resources are still in a state of growth and flux, so they will not be discussed in detail in this section. However, they are mentioned since it is likely that only an intelligent grid will be able to properly schedule a variety of distributed energy resources—and ensure that they operate safely in an interconnected grid. Intelligent Distribution Automation systems will enable a true "plug-n-play" environment, which will, in turn, truly enable the widespread use of DERs. Plug n play will also simplify system implementation for utilities.

Distribution automation systems today can also provide the means to optimize feeder and substation loading by enabling the shifting of load from one feeder to another in a very short time when needed. This same capability can yield hard dollar cost savings associated with deferment of capital projects when coupled with planning practices that take advantage of the new technologies.

Most importantly, distributed intelligence provides the tools that the utility planner will need to design a distribution system that will meet the increasing demand for reliability and power quality.

The following examples will demonstrate a wide range of system types available for IDA. Each discusses some of the benefits and drawbacks of each system and will provide a reference for the reader to consult when considering deployment of truly IDA on their system for reliability improvement, improved asset management, capital deferment, overtime reduction, improved knowledge of system conditions, and generally better customer service.

16.4.1 Automated Feeder Switching Systems

Recloser Loop Schemes. While they do not utilize distributed intelligence in the sense of intelligence shared via communication systems, recloser loop schemes are discussed because they are a fairly prevalent method for automating a system without the use of a central control logic.

Recloser-based systems typically rely on the idea that some percentage of faults on a system is temporary in nature. By reclosing some number of times for a temporary fault, sufficient time will go by for the fault to fall clear of the line and a subsequent reclosing operation will restore service. A permanent fault will not be cleared by the multiple reclosing operations and the device(s) trying to reclose will eventually lock open (lockout).

A typical line can be broken into two or perhaps three segments using multiple reclosers. The number of segments is typically limited by the ability to establish time overcurrent coordination between multiple reclosers such that only the last one before a faulted section operates to clear. A three-segment circuit would be fairly rare, as coordination would need to occur for, ultimately, four devices in series—the substation breaker, two normally closed reclosers, and, when a loop operates, a tie that would close—this typically proves quite difficult to do in practical application.

Reclosers rely on *local* overcurrent detection, voltage sensing and timers to effect restoration of the loop. When a fault occurs, the recloser immediately upstream of the fault will trip to clear the fault. It will then test the line through repeated application of fault current by reclosing a user-configured number of times. Reclosers downstream of a fault will sense loss of voltage and initiate a loss-of-voltage (LOV) timer. When the timer expires, normally open reclosers will open. A normally open tie recloser will close when its loss-of-voltage timer expires.

Recloser loop schemes typically consist of between fault interrupting reclosers, arranged in a simple loop. A 3-recloser, scheme is shown in Fig. 16-19. R1 and R3 are normally closed reclosers and R2 is a normally open tie between the two circuits.

FIGURE 16-19 3-reclosers loop scheme.

A fault between the substation SB D and R1, for example, will result in a trip of substation breaker D. It will reclose its configured number of times and will lock out for a permanent fault. R1 will sense the loss of voltage on its source and, upon expiration of its loss-of-voltage timer, will open. R2 will close upon expiration of its LOV timer and service will be restored to the unfaulted segment.

Should a permanent fault occur between R1 and R2, for example, then R1 will trip, reclose and lock out. Then, the normally open tie recloser R2 will automatically close into the fault after its loss-of-voltage timer expires in an attempt to restore service, but will trip and lock out. The unfaulted feeder emanating from substation breaker SB B will experience the fault current as well as voltage sag for all customers on the system.

Though such systems are often fitted with SCADA communication from the device back to the SCADA master station, there is no communication between devices and this type of system does not utilize distributed intelligence. Nevertheless, it is an effective way to automatically restore service to unfaulted segments and is easily implemented with little concern about communications.

But, such systems require that load capacity be reserved on each circuit to accommodate any load that may be picked up during a restoration sequence. This reserved capacity is typically based upon peak loading conditions and cannot account for actual time of day and seasonal load diversity factors. Therefore, for the bulk of the time that a circuit is not faulted, it is also an asset that is not being fully utilized.

Intelligent Loop Restoration Systems. Intelligent loop switching may use either switches or reclosers to effect automatic local sectionalizing of looped distribution circuits, then use distributed intelligence and peer-to-peer communications to effect automatic restoration of the system. A typical intelligent loop system using seven switches is shown in Fig. 16-20.

In the intelligent loop restoration system, each system switching device utilizes 3-phase voltage and current sensing to detect the passage of fault current and loss of voltage events following initiation of a fault. Each device also continuously monitors load for use in ensuring that loading limits are not exceeded during the circuit restoration process.

When a fault occurs, each device upstream of the faulted segment will see passage of fault current; each device downstream will see no fault current. All devices will see the loss-of-voltage condition when the upstream protective device operates. Logic dictates that the fault is in the line segment where the upstream switch sees fault current and the downstream switch does not.

The switching devices will open based on either counts of overcurrent or loss of voltage or upon expiration of a loss-of-voltage timer. Once this occurs, the distributed intelligence in each switch control will activate the restoration process. Based on knowledge of prefault loads in each segment and knowing the fault location, the intelligent restoration agent will close open switches only if the unfaulted segment can accommodate the load, and if the switch will not close into a faulted segment.

The use of ongoing voltage, current monitoring, and distributed intelligence ensures that the backup circuit will not be overloaded during the restoration process. This enables higher normal

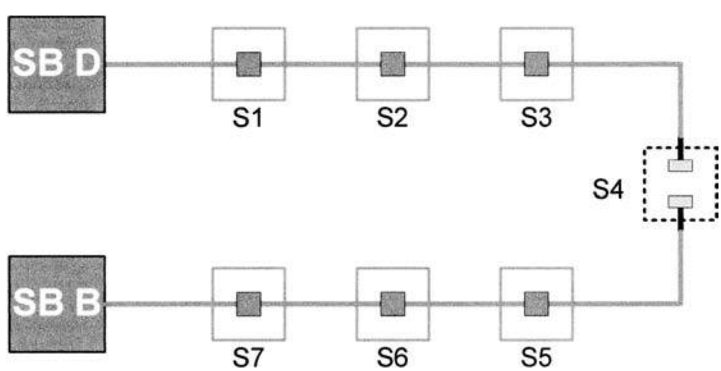


FIGURE 16-20 7-Switch intelligent loop restoration system.

loading than with noncommunicating loop schemes, which only accommodate load by reserving the capacity. Since real-time load monitoring is enabled, the system can take full advantage of load diversity, allowing restoration when able and preventing overload as needed. The result is that normal line loading can be increased to 75% of full-load capability or more depending on the amount of segmentation of the circuit.

And, distributed intelligence means that that there will be no intentional closing of a device into a faulted segment. This significantly improves power quality for customers on the backup circuit when compared to a loop scheme, which intentionally closes the backup circuit into the fault.

An intelligent loop restoration system may use reclosers if proper coordination can be achieved with the desired number of circuit segments. If this is not possible, then switches may be used and the difference will be that more customers will be affected by momentary outages than with the all-switch case. Since intelligent loop restoration schemes are designed to complete the restoration process in less than 60 s, only the customers on the faulted segment of line will see an extended outage.

A side benefit of such schemes is the reduction in line patrol time—after all, a line crew travels to the patrol site at 50 mi/h, but performs the patrol at 5 mi/h. Getting to the faulted segment faster—and reducing patrol time means that restoration is faster—and overtime is potentially reduced.

Intelligent Multigrid Switching. In order to achieve significant jumps in both reliability and greater asset utilization than the systems described above, then a system must accommodate multiple sources. In this way, it is possible for some line segments to have more than one possible alternate source. The one-line shown in Fig. 16-21 consists of four circuits with multiple possible circuit ties through normally open switches.

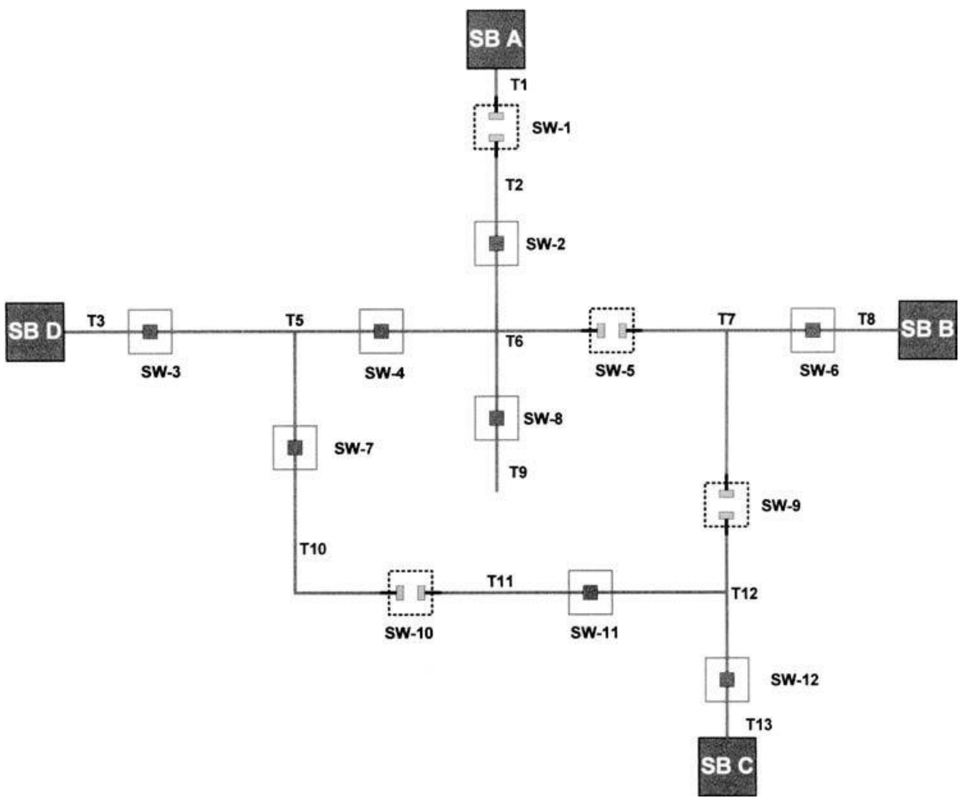


FIGURE 16-21 Intelligent multigrid system.

In the case above, all switching points utilize switches. Effecting time-overcurrent coordination between reclosers in such a system for both initial and contingency conditions would not be feasible due to the complexity of the circuit. For example, a large number of devices could wind up in series after a restoration process has been completed.

An intelligent multigrid system differs from an intelligent loop restoration system, primarily in complexity. The loop restoration system uses an overall circuit-logic approach where the entire system is treated as part of the intelligent restoration logic. The number of possible switching combinations that could come into play in a multigrid system requires a different approach.

The multigrid system breaks up the system logic such that it is resident in individual line segments bounded by intelligent switches (such as segment T6 in Fig. 16-21, which is bounded by four switches SW-4, SW-2, SW-8, and normally open SW-5). The use of a virtual agent assigned to each line segment, or team, can interact with neighboring virtual agents to effect intelligent system restoration using whatever sources are available for each de-energized segment, while still ensuring that the alternate source will not be overloaded when it is re-energized. Since the logic operates on a line segment basis, any number and type of segments can be connected to form an overall distribution system of virtually any size.

Other algorithms are used to allow for priority in choosing among multiple alternate sources when all other factors are equal. In this way, the user can force a certain amount of predictability in system operations to meet a variety of circuit planning criteria.

An example of this would be if a permanent fault occurred on line segment T5 in Fig. 16-21. In this case, substation breaker SB-D would trip, reclose, and eventually lock out. All switches on the circuit emanating from SB-D would open, either on overcurrent counts or loss-of-voltage counts. The intelligent multigrid restoration process would begin as soon as each line segment (team) confirms that the initial fault has been isolated. In the case of team T2, SW-1 will close after the agent in T2 confirms that the prefault load in T2 did not exceed the limit set for SW-1. The same process would be carried out by the agent in team T6; however, in this case, a priority has been set to restore load from team T2 first. Therefore, the process waits a predetermined time for SW-2 to become energized. The agent in team T6 then confers with T2's agent and should sufficient capacity be available to accommodate T6's prefault load, then SW-2 will close.

However, if sufficient load capacity does not exist in T2, then the agent in T6 will proceed to SW-5, where it will confer with T7's agent, perform the load analysis and close if capacity exists. Up to eight sources in a given team can be accommodated using this intelligent analysis.

If no priority is set, then the first available source with sufficient capacity will be used to restore service to a deenergized team.

The use of this distributed logic, in small, logical elements is essential for the construction of large and complex systems. Another advantage of this capability is that more than one contingency is accommodated, as long as alternate sources are available to a team for each subsequent line fault. Even if two of the sources in Fig. 16-21 were lost, some amount of load on any of the circuits could be restored using the remaining sources.

Intelligent multigrid systems do require robust communications, but the use of peer-to-peer radios or other communication devices, along with a segment-based logic will enable the restoration algorithm to function to some degree, even if some devices lose communications.

The other challenge to such a system is change. Such systems represent tools that were not available only a few years ago; conventional distribution circuit planning and design practices do not take full advantage of such systems, so a new way must be learned. However, once the ability to establish multiple circuit ties, safely and reliably, without overloading a system are incorporated into a utility design practice, then the ability to design to meet increasingly stringent customer and regulatory requirements is expanded greatly.

Intelligent Protection, Control, and Restoration Multigrid Systems. The intelligent multigrid system provides significant benefits, but requires operation of an "upstream" protective device—typically a substation breaker—to clear the initial fault. The next leap in function is to incorporate complete protection into such a system, thus eliminating outages for any but the faulted segments and beyond, for any given contingency.

Again, the use of a distributed, segment-based approach lends itself well to addressing both restoration and protection in any kind of circuit configuration. In this case, however, new algorithms and virtual agents must be used to accommodate adaptation of the protection system in addition to the comparatively straightforward restoration process outlined above.

Further, it is extremely advantageous to use highly accurate and fast protective devices and associated relays are required such that many applied in series are coordinated so that only the last device serving the faulted segment clears the fault. A method for fitting of curves based on upstream and downstream protective devices is needed—the curve-fitting agent that is responsible for this task.

This agent is also used to establish initial protective curves that provide coordination on the system in its normal state. As shown in Fig. 16-22, the curves for each substation breaker are not shown, but are designated A1, B1, C1, and D1. Protection curves for devices on the line are designed D2, D3, and D4, for example, with the higher number curve being faster than the lower numbered curve to establish a time-overcurrent coordinated system.

Once a restoration process has been completed, then the protection system must update itself to ensure that the protective coordination is maintained, regardless of the source from which a given segment is fed. An adaptive protection agent, working with the curve-fitting agent carries out this task.

In the case shown in Fig. 16-22, a loss of the source to substation SB-D will result in initiation of a restoration process, which will open the normally closed interrupting devices, then close open points, after checking for load capacity and absence of fault indication on a segment. Note that it is assumed that there is no communication between any of the field fault interrupters and the substation breaker relays.

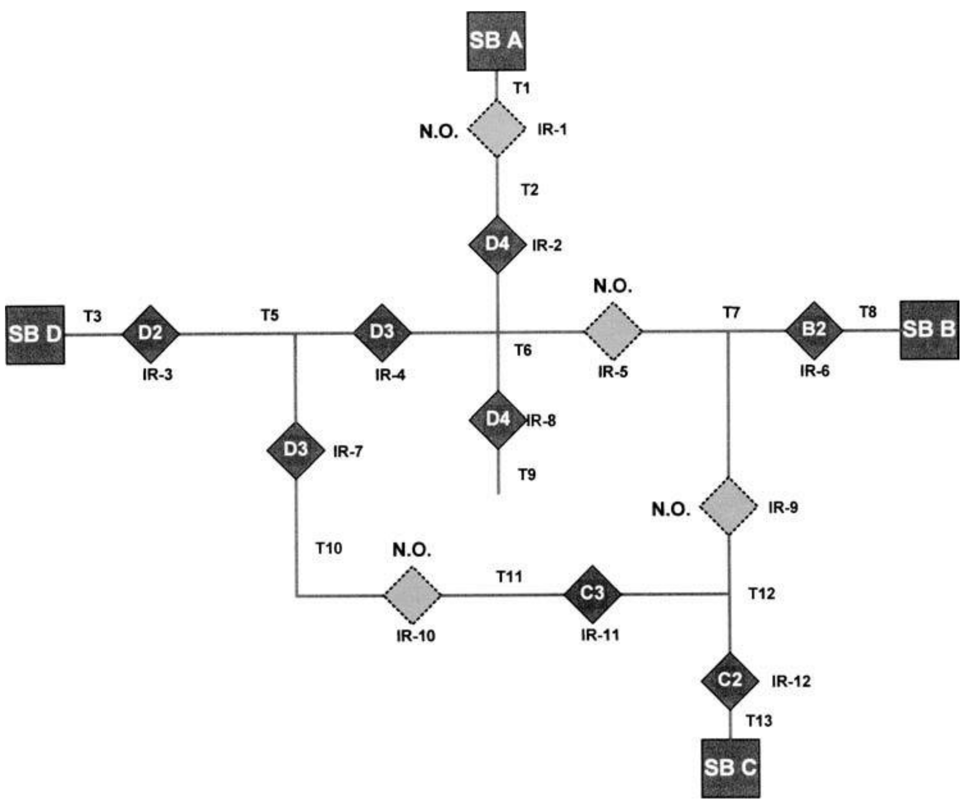


FIGURE 16-22 Intelligent protection, control and restoration multigrid system.

One possible result is for a new circuit to be configured, emanating from SB-C, feeding through IR-12, IR-11, IR-10, IR-7 and then to new open points at IR-3 and IR-4, as shown in Fig. 16-23. Using an adaptive coordination algorithm, each interrupter will talk to its upstream counterpart to ensure that it either has a faster protection setting or the same protection setting as its upstream neighbor.

It is noted that there may not always be sufficient room between protective curves for direct time-overcurrent coordination. In this case, devices must share a common curve. When this condition occurs, coordination is still possible using high-speed communications to effect coordination between devices using a blocking and adaptive coordination, process, whereby all devices that detect fault current passage signal their upstream neighbors. Interrupters that receive the signal increment their protection curve slower by 1.

For example, if a fault were to occur in segment T5 in Fig. 16-23, then interrupters IR-7, IR-10, IR-11, and IR-12 would detect the fault. Since the system uses distributed intelligence, the coordination agent knows that IR-7 and IR-10 share curves and that IR-11 has a setting of C3 which coordinates with C4. Therefore, only the interrupters with shared curves will utilize communications-based coordination. When the fault occurs, IR-7 will detect the fault and send a signal to IR-10. IR-10 will wait a small amount of time and when the signal is received, will decrement to curve C3. C4 did not receive any signal from a downstream device; therefore, it remains at C4. C4 is faster than C3 and all upstream curves; therefore, it is the only device to operate, thus clearing the fault in T5 in a coordinated fashion.

The intelligent protection, control, and restoration multigrid system combines the advantages of fault-interrupting devices while using distributed intelligence to overcome the difficulties in applying such fault-interrupting devices in complex circuit configurations. All the while, such a system also

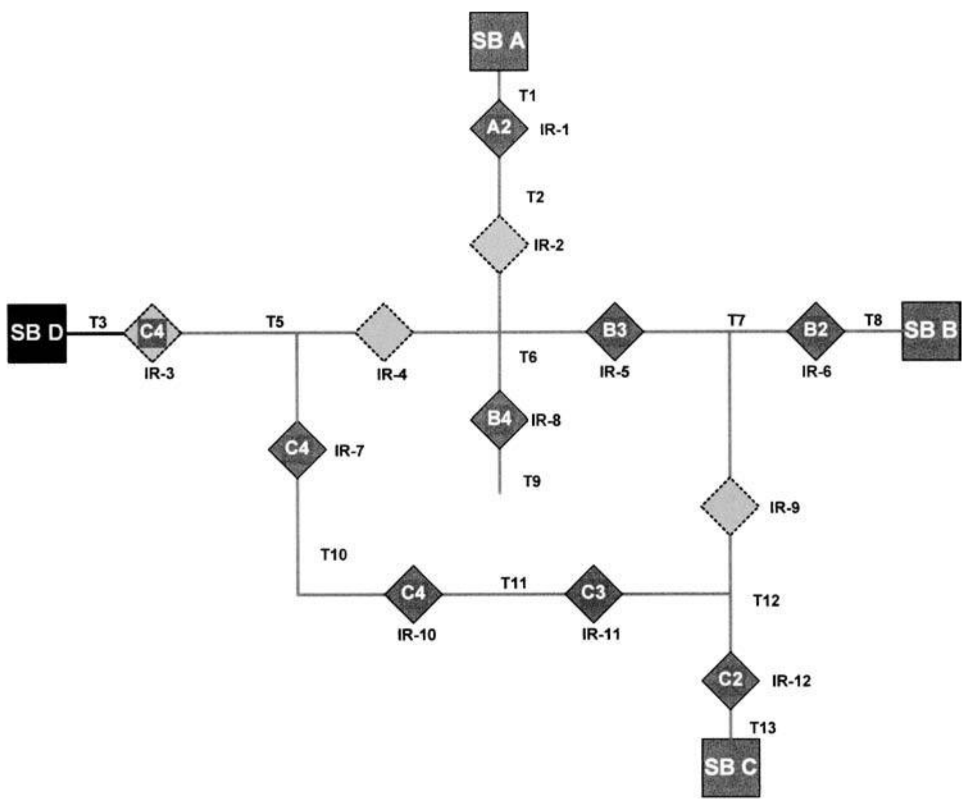


FIGURE 16-23 Reconfigured state of example system.

monitors circuit loading to prevent inadvertent overloading, thus delivering the improved asset management benefits of the intelligent multigrid switching system.

16.4.2 Summary

Each of the circuit protection and restoration systems described in this section are available in the industry as of the writing of this book. It is clear that some are easy to apply, but deliver only simple benefits. The more complex and modern a system is, the more benefits it can provide. But, it also requires a shift in thinking about circuit planning and design to maximize those benefits.

In the end, the use of highly intelligent devices on the distribution feeder—in circuit configurations that are probably difficult for many of us to even conceive of today—will be the answer to meeting the ever-increasing needs of users of electric power. Such systems will also, ultimately, reduce utility workload by moving the operating decisions to an intelligent, distributed system, taking full advantage of advances in computing and communication technologies on an ongoing basis. This will leave utility engineers and planners free to be more creative in their designs.

16.5 IMPACTS OF EFFECTIVE DSM PROGRAMS

By HESHAM SHAALAN and CHRISTA LORBER

16.5.1 Introduction

Demand-side management (DSM) is proving to be a viable means by which utilities can meet their load-shape objectives. Two decades of studies, projections, and pilot programs are suggesting that DSM can be cost-effective and flexible. Demand-side management programs have the potential to target diverse areas of end-use electricity consumption, thus deferring the need to meet growing demand through added capacity. From the utility's perspective, this promotes cash flow that would otherwise be tied up in capital investments and costs. From the customer's perspective, these programs provide incentives that range from lower electricity rates to rebates on the purchases of more efficient appliances and equipment. Therefore, there are benefits which are attractive to both parties. However, another important benefit of DSM is preserving the environment. Electricity reductions that proceed from DSM programs translate into savings by curtailing and delaying the environmental impacts for which pollutants and greenhouse gas (GHG) emissions are greatly responsible.

Commercial-sector DSM programs provide significant options for utilities in meeting growing demand. This section provides an estimation of savings in cost as well as projected GHG emissions based on effective DSM programs in the commercial sector. Realistic estimates of savings based on actual results from two previous utility studies will be presented.

Most electric utility systems in the United States were designed to account for some daily, weekly, and seasonal variability in load. This variability is desirable from the planned maintenance point of view. To account for the fluctuations that occur, different types of generating facilities are used together in various combinations to minimize total costs. This is necessary because the electric utility industry is quite capital-intensive. For every $1.00 of revenue, the utility industry requires $3.50 of capital, compared with the average industry, which needs only $0.80 per dollar of revenue.[1] Aside from the moderate fluctuations in demand, electric power is most efficiently produced when changes in the total system load are kept as small as possible. Ideally, the ratio of average power to peak power, or *load factor,* should be kept high. Interestingly, DSM provides opportunities through which utilities can achieve increasing power-system load factors.

16.5.2 Commercial-Sector DSM

Demand-side management encompasses a variety of activities that influence the pattern and magnitude of a utility's load. Programs are geared to meet one of six main objectives depending on whether the utility is targeting residential, industrial, or commercial customers, since the load curves for each of these sectors vary considerably. The objectives utilities set to change their load shape include peak clipping, valley filling, load shifting, strategic conservation, strategic load growth, and flexible load shape.

Of the preceding six objectives, peak clipping for the commercial sector is of overwhelming interest to power utilities and distribution cooperatives alike. The reason for this is historical. Using existing plants more effectively is preferable over building a new plant. For example, peak clipping makes the power system more reliable, alleviating the need for a peak-load plant that is expensive to run and only necessary 10% of the time to meet peak demand.[1] Moreover, the commercial sector has been the most rapidly growing sector in terms of electricity sales and peak. In fact, between 1970 and 1990, the sales gain for this sector averaged about 24 billion kWh per year, which is 36% of the total gain.[2] Although sales gains over the course of the next two decades are expected to be halved in comparison, electric utilities still have to contend with the increases. For these reasons, commercial-sector DSM programs provide significant options for utilities in the determination of how they will meet growing demands.

16.5.3 Effective DSM Programs and Their Impacts

For most commercial buildings, lighting and air conditioning comprise 70% of electricity consumption, where lighting accounts for 40% and air conditioning accounts for 30%.[3] Such predominance provides opportunities for significant savings that could result from well-planned programs targeting these two consumption types.

Lighting Control Program Savings. Commercial customers perceive lighting as a necessary, fixed load because inadequate or ineffective illumination hampers productivity and sales. Therefore, lighting can be classified as a predictable load from the utility's point of view. Thus, lighting control is a viable candidate for DSM programs which promote increased penetration of energy-efficient lamps and ballasts.

To exemplify the magnitude of savings that can occur, a case in point is essential. Consolidated Edison Company of New York (Con Edison), one of the largest electric utilities in the country, sponsored an Enlightened Energy Rebate Program beginning in 1991 which provided cash rebates for both retrofit and new installations of high-efficiency lighting. Five years of pilot programs precluding the Enlightened Energy Rebate Program provided the experience on which to anticipate success. Con Edison's goals were twofold in promoting the energy rebate. The first objective was to reduce the load on certain transmission and distribution (T&D) equipment, a very cost-effective goal, since increasing T&D capacity in the New York area is quite expensive. The second objective was geared toward increasing profits via incentive rates of return approved by the New York Public Service Commission. Rewards were granted for energy savings rather than capacity reduction; however, in the process of striving for the kWh savings, significant peak reductions occurred. In fact, Con Edison is projecting peak reductions of 22% to 23% by the year 2008.[4]

The Enlightened Energy Rebate Program was offered to 40,000 commercial customers, 2744 of which participated. The verified reductions in 1991 were 157.9 MW of electricity and 241 million kWh.[4] Table 16-1 translates these energy savings into emission savings in terms of quantities and associated costs. The assumptions used in calculating these values are provided in the Appendix along with sample calculations. The total cost is dominated by the CO_2 and SO_2 emission savings, which amount to \$2.35 million and \$1.39 million, respectively.

TABLE 16-1 Projected Emission and Cost Savings: Consolidated Edison Enlightened Energy Program

Millions kWh	241
CO_2 (thousand tons)	173
SO_2 (tons)	343
NO_x (tons)	187
CO (tons)	24.3
VOCs (tons)	2.76
Total cost (million dollars)	4.05
Electricity (MW)	157.9

Air-Conditioner Control Program Savings. Load control is likewise a DSM program with the potential to achieve significant penetration in the commercial sector. Currently, tens of thousands of commercial facilities have been retrofitted or were originally constructed with building energy management and control systems. Weather-sensitive loads often account for the highest peaks in demand seen by utilities, thereby degrading their annual load factors. As a result, these types of loads are excellent load-control candidates. Studies indicate that air conditioners represent the most commonly controlled commercial load.[3] This implies that improved load factors and considerable savings can be realized by air-conditioner control programs.

Available load research data and a previous air-conditioning load management program[1] sponsored by Arkansas Power and Light (AP&L) provide an opportunity to quantify savings for the commercial sector of hot springs. The previous air-conditioning load-control program took place during the summers of 1975 and 1976 and targeted two residential areas to determine the economic feasibility of interrupting central air-conditioning units for short periods of time during peak-load periods. The monitoring of this particular pilot program was implemented using a Motorola radio system with a remote-controlled switch during the hours of 1 to 5 P.M. from June 15 through September 15. Because this was a pilot program, AP&L had several objectives concerning program effectiveness. The first involved determining the contribution of a single unit to peak demand and the amount of air-conditioning load which could be displaced during the system peak-load periods. Determining the threshold of customer inconvenience incurred through implementing control during the peak periods was the second objective. The third focus of the investigation was the feasibility and reliability of a radio-controlled system.

The incentive to the participating customers included a $2.00 return per kVA of air-conditioning capacity per month, free service inspection on the air-conditioning system, and a guarantee that the system would be restored to its pretest condition should any damage result. The results from this test were exceedingly favorable. Each residential central air conditioner contributed 4 kW to the system peak and could be switched off 15 min out of each hour without causing the customer discomfort. A peak-load reduction of 1 kW per unit resulted from this control action. In addition, the tests established that radio control was a viable means of shedding loads during peak conditions. More than 25,000 residential radio switches were installed at the end of 1978, with additional installation plans of 25,000 per year until reaching the saturation goal of 125,000.

Encouraged by the air-conditioning load-control success within the residential sector, AP&L reported plans to extend air-conditioner load control to its commercial sector. In so doing, the average peak demand reduction amounted to 1.6 kW per unit.[3] Figure 16-24 shows the load curve for the month of July generated from AP&L commercial data. The dotted line represents the effect that air-conditioning load control would have if each unit were reduced by 1.6 kW. The kWh savings result from load control between 12 noon and 6:00 P.M. Due to the difference in load shape between the

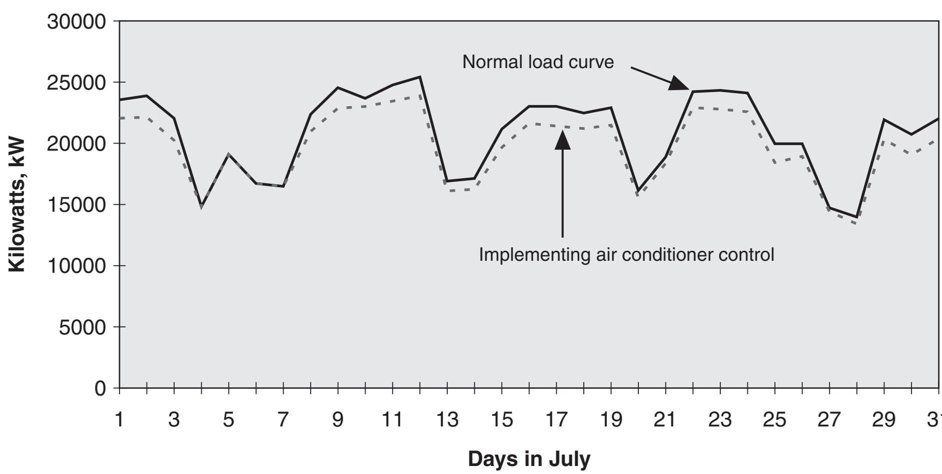


FIGURE 16-24 Load curve for the month of July.

TABLE 16-2 Projected Emission and Cost Savings for July Research Load Data: Arkansas Power & Light Commercial-Sector Air- Conditioner Control

Thousands kWh	248
CO_2 (tons)	178
SO_2 (tons)	0.353
NO_x (tons)	0.192
CO (tons)	0.025
VOCs (tons)	0.003
Total cost (dollars)	4,173
Electricity (MW)	0.248

residential and commercial sectors, the control had to be extended by 2 h in order to effectively shed peak load. Despite the additional 2 h of control, the assumption of a 15-min/h shut-off cycling period remained consistent with the residential control program.

Load data for the month of July was used in Fig. 16-24 because this is the hottest month during which AP&L must provide services. Table 16-2 quantifies the savings that result from air-conditioner control for the month of July, as seen in Fig. 16-24. Once again, CO_2 and SO_2 dominate the total cost savings. CO_2 contributes $2421 to the overall $4173, while SO_2 contributes $1433. Although the reported emissions seem relatively small, the results of Table 16-2 represent a mere 1.3% of the potential commercial customers eligible to participate in air-conditioning load control. In addition, the season for commercial-sector air conditioning is considerably longer than for the residential sector. Thus, significant emission savings are attainable with increased participation for a full season.

16.5.4 Projected Total DSM Program Impacts

Table 16-3 illustrates how all cooling and lighting programs within the commercial sector will impact the environment over the course of the next two decades. The figures reported in the table give combined totals of cooling and indoor/outdoor lighting; however, the lighting contribution is almost twice that of the cooling.

The savings in kWh, emissions, and cost reported in Table 16-3 are evidence of the effectiveness of DSM strategies. By the year 2000, CO_2 emission savings will be above 35 million tons, which alone accounts for $486 million. The electricity and added capacity savings is phenomenal, both short and long term. By 2020, savings will be more than twice that occurring in the year 2000. SO_2 and NO_x emission savings are likewise quite sizable, providing evidence that conservation and preservation can occur with DSM.

16.5.5 Conclusion

Of the many DSM programs that could enable utilities to operate efficiently, lighting and air-conditioner load control are two proven methods by which utilities can reduce their peak and provide savings for themselves, their customers, and the environment. To be effective, a program must be well planned. Therefore, successful programs are usually preceded by a pilot program to ensure

TABLE 16-3 Total Lighting and Cooling Emission and Cost Savings

	Year		
	2000	2010	2020
Billions kWh	4.98	8.50	10.9
CO_2 (million tons)	35.7	61.0	78.1
SO_2 (thousand tons)	70.8	121	155
NO_x (thousand tons)	38.6	66.0	84.4
CO (thousand tons)	5.00	8.55	10.9
VOCs (thousand tons)	0.570	0.973	1.25
Cost (million dollars)	837	1430	1830
Electricity (GW)	66.1	113	145
Added capacity (400-MW plant)	165	282	362

TABLE 16-4 Emission Characteristics of Power Plants in the United States (g/kWh)

Plant type	VOCs	CO	NO_x	SO_2	CO_2
Gas	0.025	0.20	1.00	0.004	490
Oil	0.050	0.19	1.00	5.08	781
Coal	0.010	0.11	1.00	2.00	1030

that the utility's objectives can be met. Furthermore, a utility must build a good relationship with its customer base. Otherwise, future DSM programs may be jeopardized. This can be achieved through frequent customer contact and by giving the customer some decision-making provisions. Support services necessary to a particular DSM program should be established prior to the effective starting date. This is essential in order to monitor the program accurately. Without adequate preparation, gathering data and keeping up with customer inquiries are virtually impossible.

The purpose of DSM is to improve what already exists and make what is new as efficient as possible. Meanwhile, the utility's reputation is at stake. For this reason, a major focus of these programs is the customer. In the process of promoting energy efficiency and establishing trustworthy relationships, it may be easy to overlook environmental impacts and savings. Therefore, any precautionary action that may minimize environmental changes has additional value beyond successful programs and satisfied customers. Thus, it can be argued that sustenance is the hidden value of DSM.

APPENDIX

Assumptions. The information provided herein focuses on the manner in which the values reported in Tables 16-1, 16-2, and 16-3 were calculated. Some basic assumptions have been made in order to obtain those results:

1. Coal has a 70% carbon content.
2. A 400-MW coal plant uses 800,000 MT of coal per year.
3. CO_2 recovery equipment has a 90% removal efficiency.
4. The emission characteristics of power plants in the United States (g/kWh) are shown in Table 16-4. The values shown assume a mix of 10% combustion turbines and 90% steam turbines.
5. The cost of pollutant per ton emitted is shown in Table 16-5.

TABLE 16-5 Cost of Pollutant per Ton Emitted

Pollutant	Cost ($/ton)
CO_2	13.60
SO_2	4060
NO_x	1640
CO	82
VOCs	300

Sample Calculations

- *Emission**
  Tons of pollutant:

(year kWh) × (% generated electricity) × [emission characteristic (g/kWh)] × (conversion factor)

Example	*Savings of Total SO_2 Emissions*	
Gas	Tons SO_2 = (4.98E + 10) × (0.119) × (0.004) × (1.1E − 6) =	26.10
Oil	Tons SO_2 = (4.98E + 10) × (0.039) × (5.080) × (1.1E − 6) =	10,866.76
Coal	Tons SO_2 = (4.98E + 10) × (0.546) × (2.000) × (1.1E − 6) =	59,895.52
	Tons total	70,788.38

*Note that this calculation is performed for all substances listed in the preceding table.

- *Cost*
  Total cost

$$(\text{Total emissions}) \times (\text{cost of pollutant})$$

Example *Total SO_2 Cost Savings*
Cost (million \$) = (70,788) × (4060) = 287

- *Electricity and added capacity savings (based on CO_2 emissions)*
  Added Capacity Savings

$$(CO_2 \text{ emission savings})/(\text{tons of } CO_2 \text{ emitted from 400-MW plant})$$

Electricity (that could be produced with neutral GHG effect)

$$(\text{Added capacity savings}) \times [\text{size of plant (MW)}]$$

Example

Tons of CO_2 emissions = (800,000) × (2.57) × (0.1) = 216,000
Added capacity savings (No. of 400-MW plants − year 2000)
= (35,700,000)/(216,000) = 165
Electricity (GW) = (165) × (400) = 66.1

REFERENCES

1. Chamberlin, J. H., and Faruqui, A.: "Demand-Side Management: The Next Generation." Knoxville, Tenn.: Forum for Applied Research, Barakat & Chamberlin, Inc., September 30, 1991.
2. Demand-Side Management, "Drivers of Electricity Growth and the Role of Utility Demand-Side Management," Electric Power Research Institute (EPRI), Report TR-102639, August 1993.
3. Demand-Side Management, "Impact of Demand-Side Management on Future Customer Electricity Demand: An Update," Electric Power Research Institute (EPRI), Report CU-6953, September 1990.
4. Demand-Side Management, "Lessons Learned in Commercial Sector Demand-Side Management," Electric Power Research Institute (EPRI), Report TR-102551, October 1993.
5. Demand-Side Management, "1987 Survey of Commercial-Sector Demand-Side Management Programs," Electric Power Research Institute (EPRI), Report CU-6294, March 1989.
6. Talukdar, S. N., and Gellings, C. W.: *Load Management*. New York: IEEE Press, 1987.

SECTION 17
SUBSTATIONS

W. Bruce Dietzman
Project Engineering Manager, TXU Electric Delivery Company; Senior Member, IEEE; Past Chair, IEEE/PES Substations Committee; Past VP-Technical Activities, IEEE/PES

Philip C. Bolin
General Manager, Substation Division Mitsubishi Electric Power Products, Inc.; Fellow, IEEE; Past Chair, IEEE/PES Gas-Insulated Substations Subcommittee; Member, CIGRE Working Group 23.10 GIS

CONTENTS

17.1 AIR-INSULATED SUBSTATIONS

17.1.1 Function of Substations

Transmission and Distribution Systems. In large, modern ac power systems, the transmission and distribution systems function to deliver bulk power from generating sources to users at the load centers. Transmission systems generally include generation switchyards, interconnecting transmission lines, autotransformers, switching stations, and step-down transformers. Distribution systems include primary distribution lines or networks, transformer banks, and secondary lines or networks, all of which serve the load area.

17.1.2 Design Objectives

As an integral part of the transmission or distribution systems, the substation or switching station functions as a connection and switching point for generation sources, transmission or subtransmission lines, distribution feeders, and step-up and step-down transformers. The design objective for the substation is to provide as high a level of reliability and flexibility as possible while satisfying system requirements and minimizing total investment costs.

Voltage Levels. The selection of optimal system voltage levels depends on the load to be served and the distance between the generation source and the load. Many large power plants are located great distances from the load centers to address energy sources or fuel supplies, cooling methods, site costs and availability, and environmental concerns. For these reasons, the use of transmission voltages as high as 765 kV has occurred. Transmission system substations that provide bulk power operate at voltages from 69 to 765 kV. Common voltage classes used in the United States for major substations include 69, 115, 138, 161, and 230 kV (considered *high voltage* or *HV class*) and 345,500, and 765 kV (considered *extra high voltage* or *EHV class*). Even higher voltages which include 1100 and 1500 kV have been considered. These are referred to as *ultra high voltage* or *UHV class*. Distribution system substations operate at secondary voltage levels from 4 to 69 kV.

Design Considerations. Many factors influence the selection of the proper type of substation for a given application. This selection depends on such factors as voltage level, load capacity, environmental considerations, site space limitations, and transmission-line right-of-way requirements. While also considering the cost of equipment, labor, and land, every effort must be made to select a substation type that will satisfy all requirements at minimum costs. The major substation costs are reflected in the number of power transformers, circuit breakers, and disconnecting switches and their associated structures and foundations. Therefore, the bus layout and switching arrangement selected will determine the number of the devices that are required and in turn the overall cost. The choice of insulation levels and coordination practices also affects cost, especially at EHV. A drop of one level in basic insulation level (BIL) can reduce the cost of major electrical equipment by thousands of dollars. A careful analysis of alternative switching schemes is essential and can result in considerable savings by choosing the minimum equipment necessary to satisfy system requirements.

A number of factors must be considered in the selection of bus layouts and switching arrangements for a substation to meet system and station requirements. A substation must be safe, reliable, economical, and as simple in design as possible. The design also should provide for further expansion, flexibility of operation, and low maintenance costs.

The physical orientation of the transmission-line routes often dictates the substation's location, orientation, and bus arrangement. This requires that the selected site allow for a convenient arrangement of the lines to be accomplished.

For reliability, the substation design should reduce the probability of a total substation outage caused by faults or equipment failure and should permit rapid restoration of service after a fault or failure occurs. The layout also should consider how future additions and extensions can be accomplished without interrupting service.

Bus Schemes. The substation design or scheme selected determines the electrical and physical arrangement of the switching equipment. Different bus schemes can be selected as emphasis is shifted between the factors of safety, reliability, economy, and simplicity dictated by the function and importance of the substation.

The substation bus schemes used most often are

1. Single bus
2. Main and transfer bus
3. Double bus, single breaker
4. Double bus, double breaker
5. Ring bus
6. Breaker and a half

Some of these schemes may be modified by the addition of bus-tie breakers, bus sectionalizing devices, breaker bypass facilities, and extra transfer buses. Figures 17-1 to 17-6 show one-line diagrams for some of the typical schemes listed above.

Single Bus. The single-bus scheme (Fig. 17-1) is not normally used for major substations. Dependence on one main bus can cause a serious outage in the event of breaker or bus failure without the use of mobile equipment. The station must be deenergized in order to carry out bus maintenance or add bus extensions. Although the protective relaying is relatively simple for this scheme, the single-bus scheme is considered inflexible and subject to complete outages of extended duration.

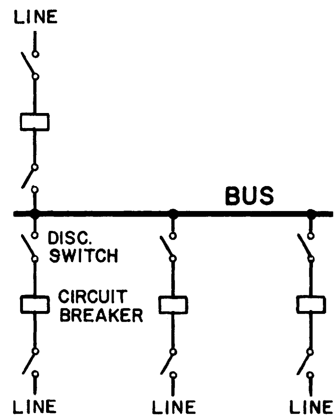


FIGURE 17-1 Single bus.

Main and Transfer Bus. The main- and transfer-bus scheme (Fig. 17-2) adds a transfer bus to the single-bus scheme. An extra bus-tie circuit breaker is provided to tie the main and transfer buses together.

When a circuit breaker is removed from service for maintenance, the bus-tie circuit breaker is used to keep that circuit energized. Unless the protective relays are also transferred, the bus-tie relaying must be capable of protecting transmission lines or generation sources. This is considered rather unsatisfactory because relaying selectivity is poor.

A satisfactory alternative consists of connecting the line and bus relaying to current transformers located on the lines rather than on the breakers. For this arrangement, line and bus relaying need not be transferred when a circuit breaker is taken out of service for maintenance, with the bus-tie breaker used to keep the circuit energized.

If the main bus is ever taken out of service for maintenance, no circuit breakers remain to protect any of the feeder circuits. Failure of any breaker or failure of the main bus can cause complete loss of service of the station.

Due to its relative complexity, disconnect-switch operation with the main- and transfer-bus scheme can lead to operator error and a possible outage. Although this scheme is low in cost and enjoys some popularity, it may not provide as high a degree of reliability and flexibility as required.

Double Bus, Single Breaker. This scheme uses two main buses, and each circuit includes two bus selector disconnect switches. A bus-tie circuit (Fig. 17-3) connects to the two main buses and, when closed, allows transfer of a feeder from one bus to the other bus without deenergizing the feeder circuit by operating the bus selector disconnect switches. The circuits may all operate from either the no. 1 or no. 2 main bus, or half the circuits may be operated off either bus. In the first case, the station

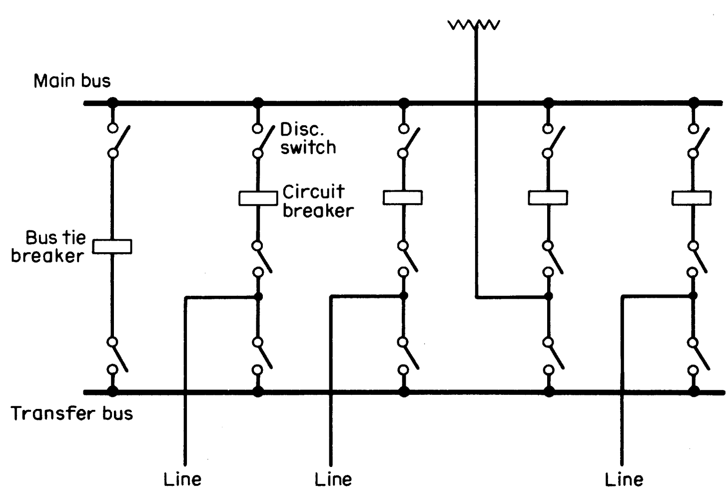


FIGURE 17-2 Main and transfer bus.

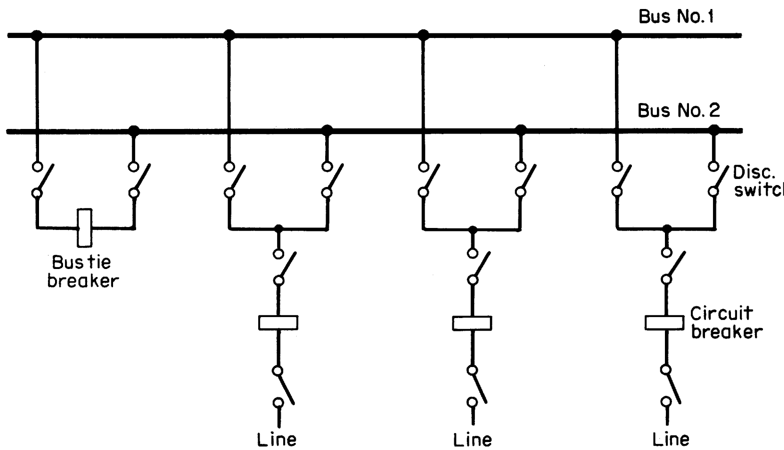


FIGURE 17-3 Double bus, single breaker.

will be out of service for bus or breaker failure. In the second case, half the circuits will be lost for bus or breaker failure.

In some cases circuits operate from both the no. 1 and no. 2 bus, and the bus-tie breaker is normally operated closed. For this type of operation, a very selective bus-protective relaying scheme is required to prevent complete loss of the station for a fault on either bus. Disconnect-switch operation becomes quite involved, with the possibility of operator error, injury, and possible outage. The double-bus, single-breaker scheme is relatively poor in reliability and is not normally used for important substations.

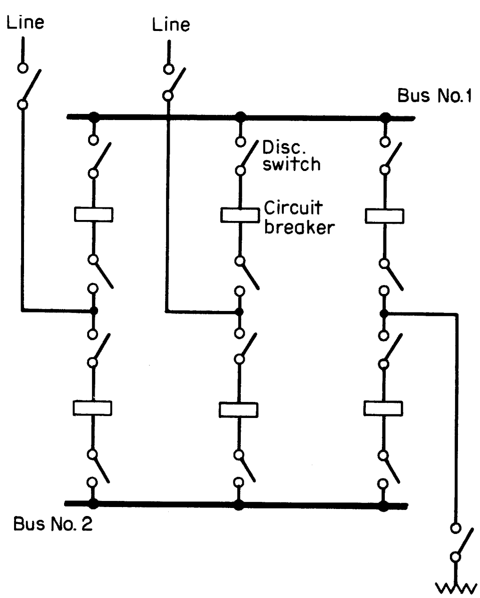


FIGURE 17-4 Double bus, double breaker.

Double Bus, Double Breaker. The double-bus, double breaker scheme (Fig. 17-4) requires two circuit breakers for each feeder circuit. Normally, each circuit is connected to both buses. In some cases, half the circuits operate on each bus. For these cases, a bus or breaker failure would cause loss of only half the circuits, which could be rapidly corrected through switching. The physical location of the two main buses must be selected in relation to each other to minimize the possibility of faults spreading to both buses. The use of two breakers per circuit makes this scheme expensive; however, it does represent a high degree of reliability.

Ring Bus. In the ring-bus scheme (Fig. 17-5), the breakers are arranged in a ring with circuits connected between breakers. There are the same number of circuits as there are breakers. During normal operation, all breakers are closed. For a circuit fault, two breakers are tripped, and in the event that one of the breakers fails to operate to clear the fault, an additional circuit will be tripped by operation of breaker-failure backup relays. During breaker maintenance, the ring is broken, but all lines remain in service.

The circuits connected to the ring are arranged so that sources are alternated with loads. For an extended circuit outage, the line-disconnect switch may be opened, and the ring can be closed. No changes to protective relays are required for any of the various operating conditions or during maintenance.

The ring-bus scheme is relatively economical in cost, has good reliability, is flexible, and is normally considered suitable for important substations up to a limit of five circuits. Protective relaying and automatic reclosing are more complex than for previously described schemes. It is common practice to build major substations initially as a ring bus; for more than five outgoing circuits, the ring bus is usually converted to the breaker-and-a-half scheme.

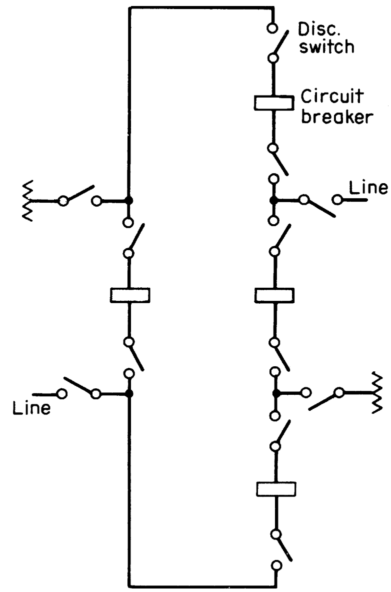


FIGURE 17-5 Ring bus.

Breaker and a Half. The breaker-and-a half scheme (Fig. 17-6), sometimes called the *three-switch scheme*, has three breakers in series between two main buses. Two circuits are connected between the three breakers, hence the term *breaker and a half*. This pattern is repeated along the main buses so that one and a half breakers are used for each circuit.

Under normal operating conditions, all breakers are closed, and both buses are energized. A circuit is tripped by opening the two associated circuit breakers. Tie-breaker failure will trip one additional circuit, but no additional circuit is lost if a line trip involves failure of a bus breaker. Either bus may be taken out of service at any time with no loss of service. With sources connected opposite to loads, it is possible to operate with both buses out of service. Breaker maintenance can be done with no loss of service, no relay changes, and simple operation of the breaker disconnects.

The breaker-and-a-half arrangement is more expensive than the other schemes, with the exception of the double-breaker, double-bus scheme, and protective relaying and automatic reclosing schemes are more complex than for other schemes. However, the breaker-and-a half scheme is superior in flexibility, reliability, and safety.

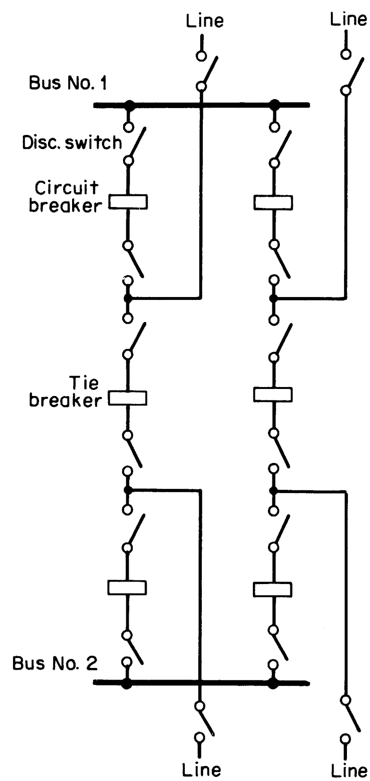


FIGURE 17-6 Breaker-and-a-half scheme.

17.1.3 Reliability Comparisons

The various schemes have been compared to emphasize their advantages and disadvantages. The basis of comparison to be employed is the economic justification of a particular degree of reliability. The determination of the degree of reliability involves an appraisal of anticipated operating conditions and the continuity of service required by the load to be served. Table 17-1 contains a summary of the comparison of switching schemes to show advantages and disadvantages.

TABLE 17-1 Summary of Comparison of Switching Schemes

Switching scheme	Advantages	Disadvantages
1. Single bus	1. Lowest cost.	1. Failure of bus or any circuit breaker results in shutdown of entire substation. 2. Difficult to do any maintenance. 3. Bus cannot be extended without completely deenergizing substation. 4. Can be used only where loads can be interrupted or have other supply arrangements.
2. Double bus, double breaker	1. Each circuit has two dedicated breakers. 2. Has flexibility in permitting feeder circuits to be connected to either bus. 3. Any breaker can be taken out of service for maintenance. 4. High reliability.	1. Most expensive. 2. Would lose half of the circuits for breaker failure if circuits are not connected to both buses.
3. Main and transfer	1. Low initial and ultimate cost. 2. Any breaker can be taken out of service for maintenance. 3. Potential devices may be used on the main bus for relaying.	1. Requires one extra breaker for the bus tie. 2. Switching is somewhat complicated when maintaining a breaker. 3. Failure of bus or any circuit breaker results in shutdown of entire substation.
4. Double bus, single breaker	1. Permits some flexibility with two operating buses. 2. Either main bus may be isolated for maintenance. 3. Circuit can be transferred readily from one bus to the other by use of bus-tie breaker and bus selector disconnect switches.	1. One extra breaker is required for the bus tie. 2. Four switches are required per circuit. 3. Bus protection scheme may cause loss of substation when it operates if all circuits are connected to that bus. 4. High exposure to bus faults. 5. Line breaker failure takes all circuits connected to that bus out of service. 6. Bus-tie breaker failure takes entire substation out of service.
5. Ring bus	1. Low initial and ultimate cost. 2. Flexible operation for breaker maintenance. 3. Any breaker can be removed for maintenance without interrupting load. 4. Requires only one breaker per circuit. 5. Does not use main bus. 6. Each circuit is fed by two breakers. 7. All switching is done with breakers.	1. If a fault occurs during a breaker maintenance period, the ring can be separated into two sections. 2. Automatic reclosing and protective relaying circuitry rather complex. 3. If a single set of relays is used, the circuit must be taken out of service to maintain the relays. (Common on all schemes.) 4. Requires potential devices on all circuits since there is no definite potential reference point. These devices may be required in all cases for synchronizing, live line, or voltage indication. 5. Breaker failure during a fault on one of the circuits causes loss of one additional circuit owing to operation of breaker-failure relaying.
6. Breaker and a half	1. Most flexible operation. 2. High reliability. 3. Breaker failure of bus side breakers removes only one circuit from service. 4. All switching is done with breakers. 5. Simple operation; no disconnect switching required for normal operation. 6. Either main bus can be taken out of service at any time for maintenance. 7. Bus failure does not remove any feeder circuits from service.	1. $1\frac{1}{2}$ breakers per circuit. 2. Relaying and automatic reclosing are somewhat involved since the middle breaker must be responsive to either of its associated circuits.

17.1.4 Arrangements and Equipment

Once a determination of the switching scheme best suited for a particular substation application is made, it is necessary to consider the station arrangement and equipment that will satisfy the many physical requirements of the design. Available to the design engineer are the following:

1. Conventional outdoor air-insulated open-type bus-and-switch arrangement substations (using either a strain bus or rigid bus design)
2. Metal-clad or metal-enclosed substations
3. Gas (sulfur hexafluoride)–insulated substations

Outdoor open-type bus-and-switch arrangements generally are used because of their lower cost, but they are larger in overall physical size. Metal-clad substations generally are limited to 38 kV. Gas-insulated substations are generally the highest in cost but smallest in size.

Substation Components. The electrical equipment in a typical substation can include the following:

Circuit breakers
Disconnecting switches
Grounding switches
Current transformers
Voltage transformers or capacitor voltage transformers
Coupling capacitors
Line traps
Surge arresters
Power transformers
Shunt reactors
Current-limiting reactors
Station buses and insulators
Grounding systems
Series capacitors
Shunt capacitors

Support Structures. In order to properly support, mount, and install the electrical equipment, structures made of steel, aluminum, wood, or concrete and associate foundations are required. The typical open-type substation requires strain structures to support the transmission-line conductors; support structures for disconnecting switches, current transformers, potential transformers, lightning arresters, and line traps, capacitor voltage transformers; and structures and supports for the strain and rigid buses in the station.

When the structures are made of steel or aluminum, they require concrete foundations; however, when they are made of wood or concrete, concrete foundations are not required. Additional work is required to design concrete foundations for supporting circuit breakers, reactors, transformers, capacitors, and any other heavy electrical equipment.

Substation-equipment support structures fabricated of steel or aluminum may consist of single wide-flange or tubular-type columns, rigid-frame structures composed of wide flanges or tubular sections, or lattice structures composed of angle members. Substation strain structures can be wood or concrete pole structures, aluminum or steel lattice-type structures, or steel A-frame structures. Aluminum, weathering steel, and concrete pole structures can be used in their natural unfinished state. Normal carbon-steel structures should have galvanized or painted finishes. Wood structures should have a thermal- or pressure-process-applied preservative finish.

Aluminum structures are lightweight, have an excellent strength-to-weight ratio, and require little maintenance but have a greater initial cost than steel structures. Weathering-steel structures can be field-welded without the special surface preparation and touch-up work required on galvanized or painted steel structures, and the self-forming protective corrosion oxide eliminates maintenance. In addition, the weathering-steel color blends well in natural surroundings. Galvanized- or painted-steel structures have a slightly lower initial cost than weathering-steel structures; however, they require special treatment before and after field welding and require more maintenance.

Lattice-type structures are light in weight, have a small wind-load area, and are low in cost. Single-column support structures and rigid-frame structures require little maintenance, are more aesthetically pleasing, and can be inspected more quickly than lattice structures, but they have a greater initial cost. In order to reduce erection costs, rigid-frame structures should be designed with bolted field connections.

The design of supporting structures is affected by the phase spacings and ground clearances required, by the types of insulators, by the length and weight of buses and other equipment, and by wind and ice loading. For data on wind and ice loadings, see National Electric Safety Code©, IEEE Standard C2-2002, or latest edition. For required clearances and phase spacings, see Part I, Secs. 11 and 12.

Other structural and concrete work required in the substation includes site selection and preparation, roads, control houses, manholes, conduits, ducts, drainage facilities, catch basins, oil containment, and fences.

17.1.5 Site Selection

Civil engineering work associated with the substation should be initiated as early as possible in order to ensure that the best available site is selected. This work includes a study of the topography and drainage patterns of the area together with a subsurface soil investigation. The information obtained from the subsurface soil investigation also will be used to determine the design of the substation foundations. For large substations or substations located in area with poor soils, it may be necessary to obtain additional subsurface soil tests after final selection of the substation site has been made. The additional information should fully describe the quality of the soil at the site, since the data will be used to design equipment foundations.

Open-Bus Arrangement. An air-insulated, open-bus substation arrangement consists essentially of open-bus construction using either rigid- or strain-bus design such as the breaker-and-a-half arrangement shown in Fig. 17-7; the buses are arranged to run the length of the station and are located toward the outside of the station. The transmission-line exits cross over the main bus and are dead-ended on takeoff tower structures. The line drops into the bay in the station and connects to the disconnecting switches and circuit breakers.

Use of this arrangement requires three distinct levels of bus to make the necessary crossovers and connections to each substation bay. Typical dimensions of these levels at 230 kV are 16 ft for the first level above ground, 30 ft high for the main bus location, and 57 ft for the highest level of bus (see Fig. 17-7).

This arrangement, in use since the mid-1920s and widely used by many electric utilities, has the advantage of requiring a minimum of land area per bay and relative ease of maintenance, and it is ideally suited to a transmission-line through-connection where a substation must be inserted into a transmission line.

Inverted Bus. An alternate arrangement is the inverted-bus, breaker-and-a-half scheme for EHV substations. A typical layout is outlined in Fig. 17-8. A one-line diagram of a station showing many variations of the inverted-bus scheme is presented in Fig. 17-9. With this arrangement, all outgoing circuit takeoff towers are located in the outer perimeter of the substation, eliminating the crossover of line or exit facilities. Main buses are located in the middle of the substation, with all disconnecting switches, circuit breakers, and bay equipment located outboard of the main buses. The end result of

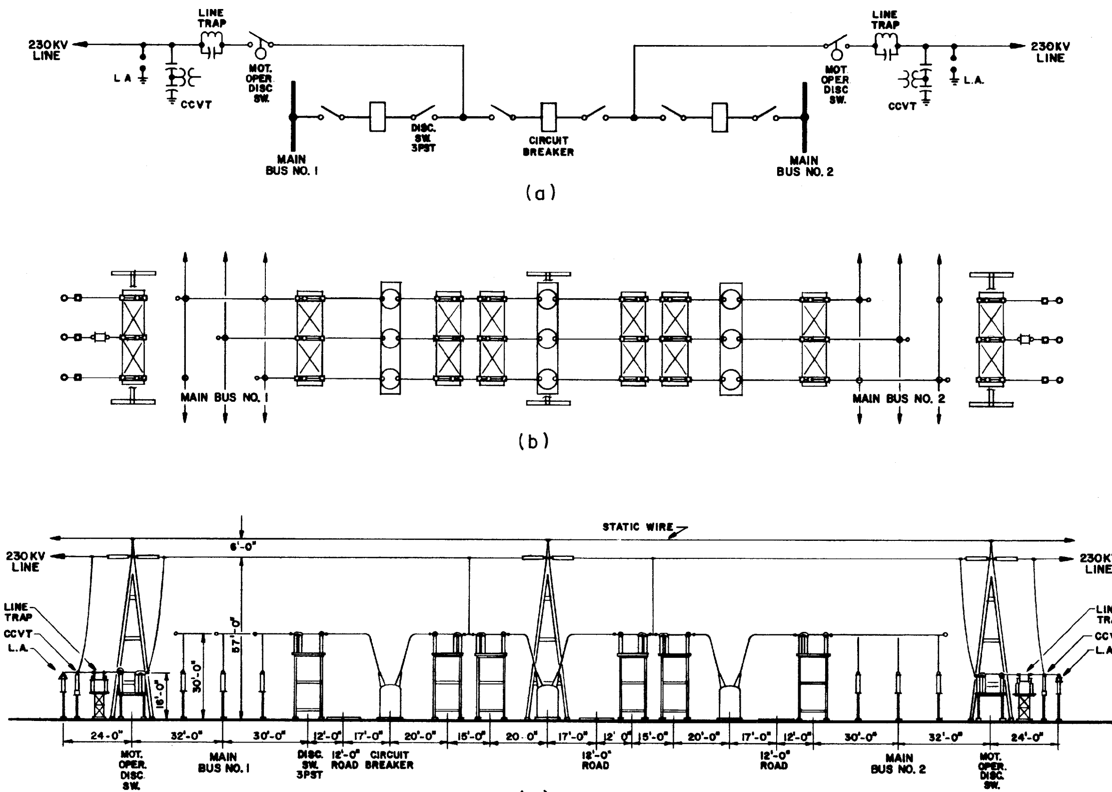


FIGURE 17-7 Typical conventional substation layout, breaker-and-a-half scheme. (*a*) Main one-line diagram; (*b*) plan; (*c*) elevation.

the inverted-bus arrangement presents a very low profile station with many advantages in areas where beauty and aesthetic qualities are a necessity for good public relations. The overall height of the highest bus in the 230-kV station just indicated reduces from a height of 57 ft above ground in the conventional arrangement to a height of only 30 ft above ground for the inverted-bus low-profile scheme.

17.1.6 Substation Buses

Substation buses are an important part of the substation because they carry electric currents in a confined space. They must be carefully designed to have sufficient structural strength to withstand the maximum stresses that may be imposed on the conductors, and in turn on the supporting structures, due to short-circuit currents, high winds, and ice loadings.

During their early development, HV class substations were usually of the strain-bus design. The strain bus is similar to a transmission line and consists of a conductor such as ACSR (aluminum cable steel reinforced), copper, or high-strength aluminum alloy strung between substation structures. EHV substations normally use the rigid-bus approach and enjoy the advantage of low station profile and ease of maintenance and operation (see Fig. 17-8). The mixing of rigid- and strain-bus construction is normally employed in the conventional arrangement shown in Fig. 17-7. Here, the main buses use rigid-bus design, and the upper buses between transmission towers are of strain-bus design. A typical design at 765 kV uses a combination of both rigid and strain buses (Fig. 17-10).

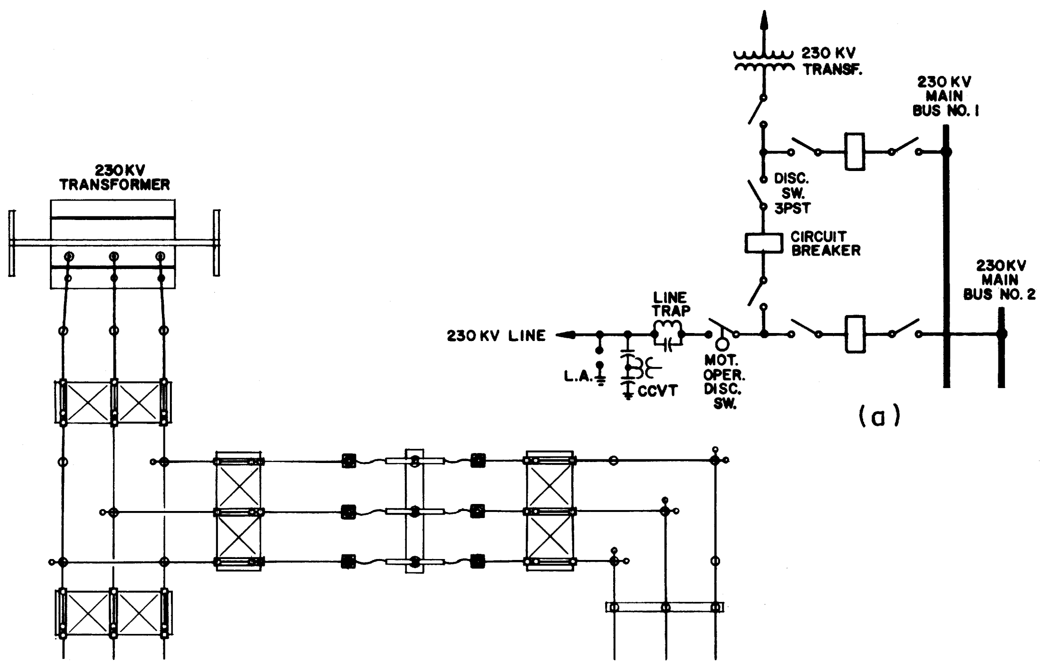


FIGURE 17-8 Typical 130-kV inverted-bus substation. (*a*) One-line diagram; (*b*) plan; (*c*) elevation.

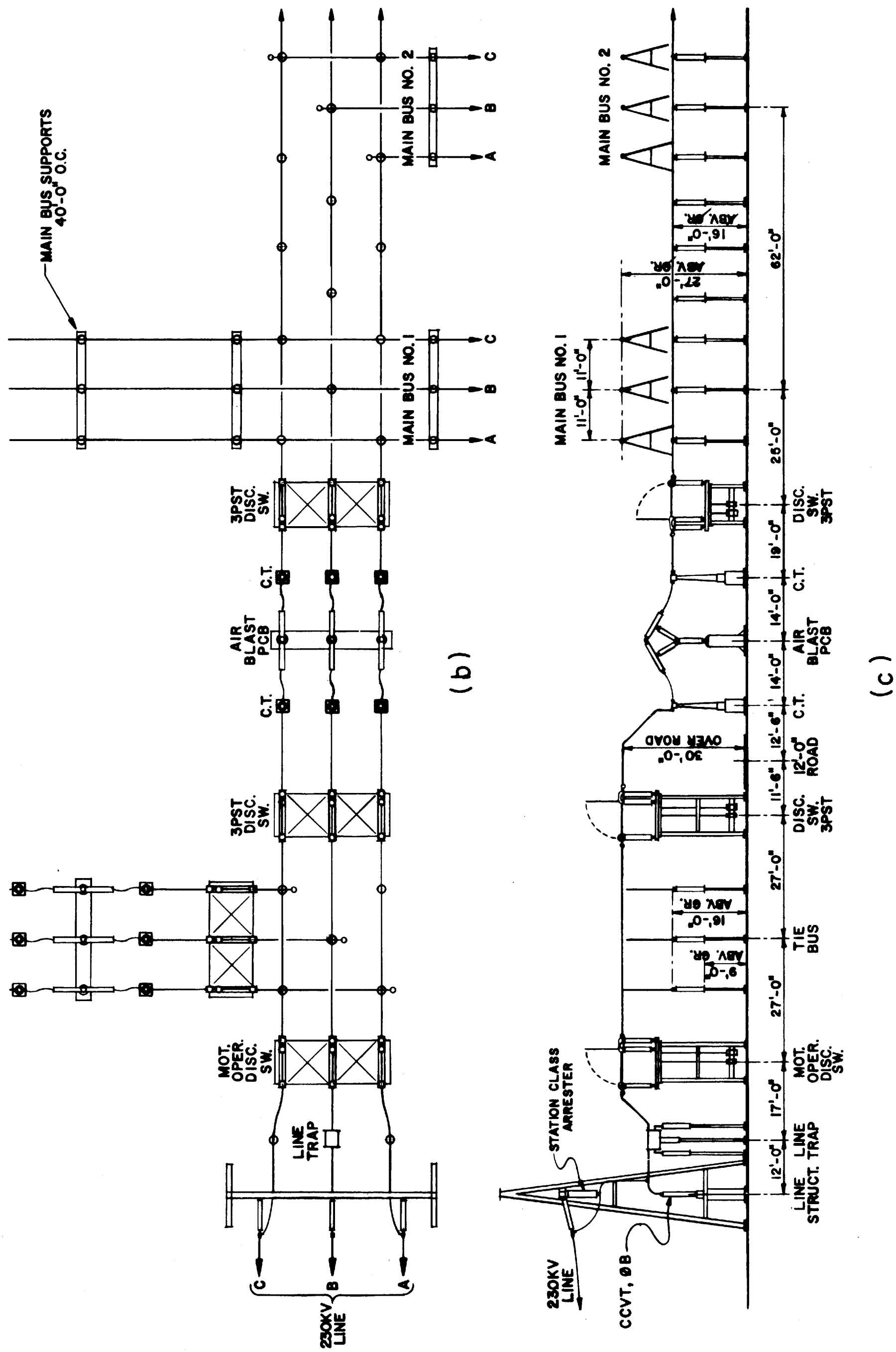


FIGURE 17-8 Typical 130-kV inverted-bus substation. (*a*) One-line diagram; (*b*) plan; (*c*) elevation.

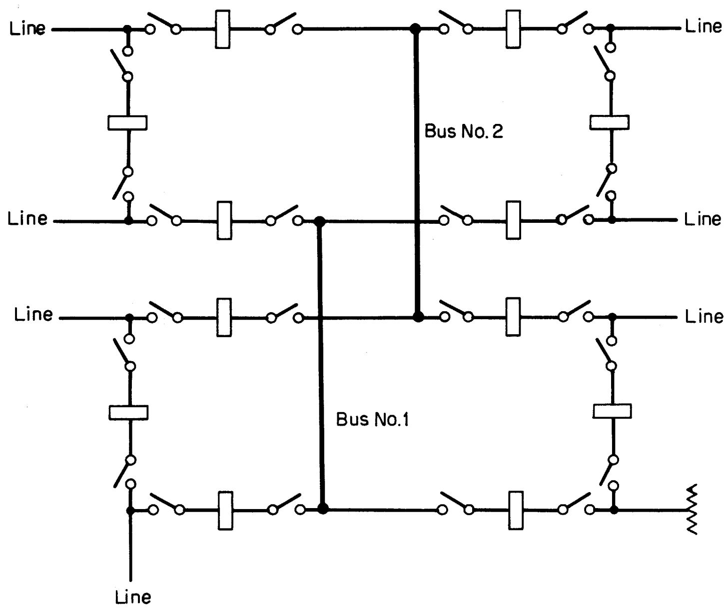


FIGURE 17-9 EHV substation, low-profile, inverted breaker-and-a-half scheme.

A comparison of rigid and strain buses indicates that careful consideration should be given to selection of the proper type of bus to use.

Rigid-bus advantages:

1. Less steel is used, and structures are of a simpler design.
2. Rigid conductors are not under constant strain.
3. Individual pedestal-mounted insulators are more accessible for cleaning.
4. The rigid bus is lower in height, has a distinct layout, and can be definitely segregated for maintenance.
5. Low profile with the rigid bus provides good visibility of the conductors and apparatus and gives a good appearance to the substation.

Rigid-bus disadvantages:

1. More insulators and supports are usually needed for rigid-bus design, thus requiring more insulators to clean.
2. The rigid bus is more sensitive to structural deflections, causing misalignment problems and possible damage to the bus.
3. The rigid bus usually requires more land area than the strain bus.
4. Rigid-bus designs are comparatively expensive.

Strain-bus advantages:

1. Comparatively lower cost than the rigid bus.
2. Substations employing the strain bus may occupy less land area than stations using the rigid bus.
3. Fewer structures are required.

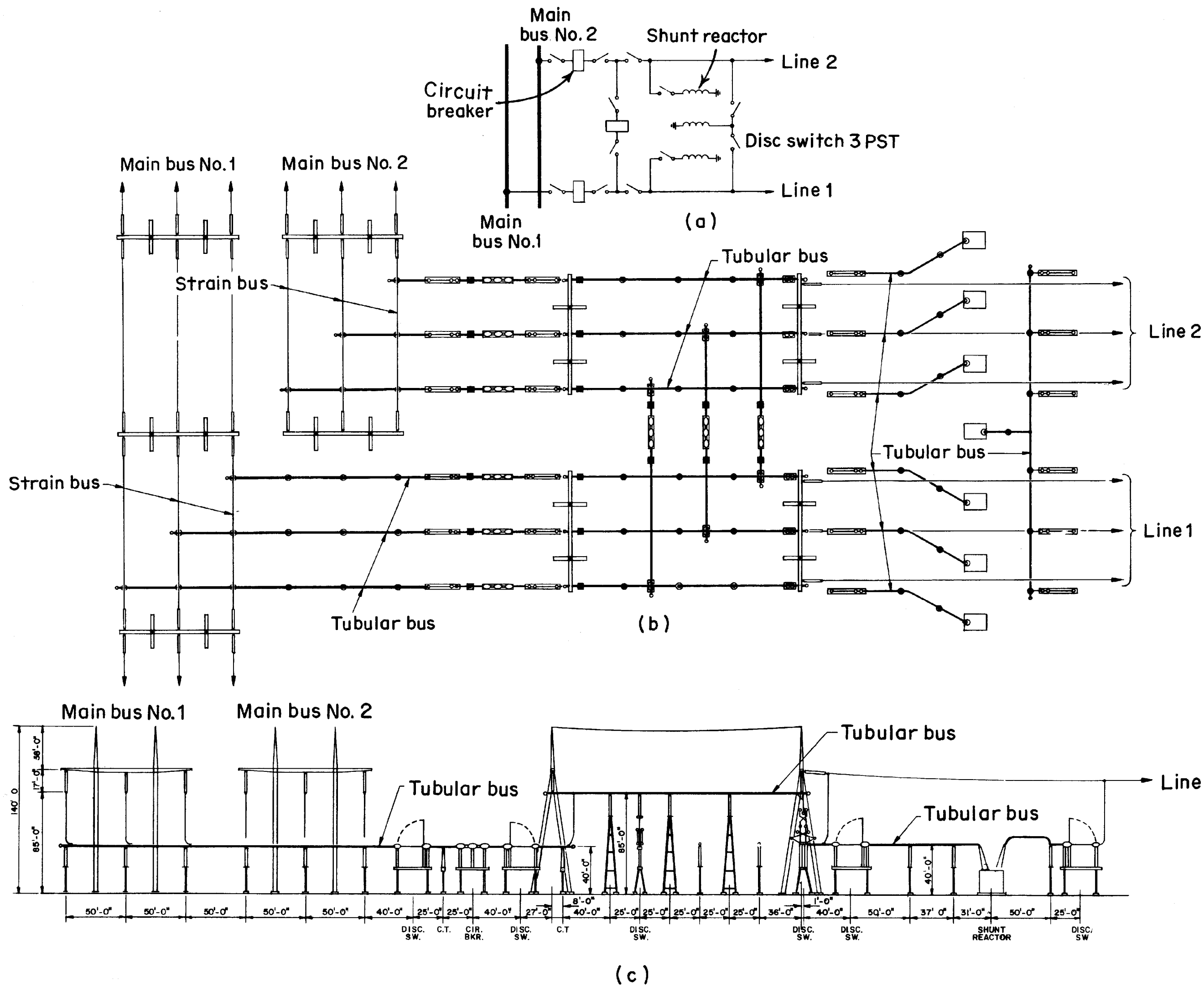


FIGURE 17-10 A 765-kV substation using both rigid- and strain-bus design. (*a*) Main one-line diagram; (*b*) plan; (*c*) elevation.

Strain-bus disadvantages:

1. Strain structures require larger structures and foundations.
2. Insulators are not conveniently accessible for cleaning.
3. Painting of high-steel structures is costly and hazardous.
4. Emergency conductor repairs are more difficult.

The design of station buses depends on a number of elements, which include the following:

1. Current-carrying capacity
2. Short-circuit stresses
3. Minimum electrical clearances

The current-carrying capacity of a bus is limited by the heating effects produced by the current. Buses generally are rated on the basis of the temperature rise, which can be permitted without danger of overheating equipment terminals, bus connections, and joints.

The permissible temperature rise for plain copper and aluminum buses is usually limited to 30°C above an ambient temperature of 40°C. This value is the accepted standard of IEEE, NEMA, and ANSI. This is an average temperature rise; a maximum or hot-spot temperature rise of 35°C is permissible. Many factors enter into the heating of a bus, such as the type of material used, the size and shape of the conductor, the surface area of the conductor and its condition, skin effect, proximity effect, conductor reactance, ventilation, and inductive heating caused by the proximity of magnetic materials.

Rigid-Bus Material. Rigid-bus materials in general use are aluminum and copper. Hard-drawn aluminum, especially in the tubular shape, is the most widely used material in HV and EHV open-type outdoor stations. Aluminum has the advantage of being about one-third the weight of copper and requires little maintenance. The proper use of alloys of aluminum will provide the rigidity needed to serve as a bus material. For a given current rating and for equal limiting temperatures, the required area of aluminum bus is about 133% of the area of the copper bus. Copper and aluminum tubing, as well as other special shapes, are also used for low-voltage distribution substation buses.

Skin Effect. *Skin effect* in a conductor carrying an alternating current is the tendency toward crowding of the current into the outer layer, or "skin," of the conductor due to the self-inductance of the conductor. This results in an increase in the effective resistance of the conductor and in a lower current rating for a given temperature rise. Skin effect is very important in heavy-current buses where a number of conductors are used in parallel, because it affects not only each conductor but also each group of conductors as a unit.

Tubing has less skin-effect resistance than rod or flat conductors of the same cross section, and tubing with a thin wall is affected the least by skin effect. Aluminum conductors are affected less by skin effect than copper conductors of similar cross section because of the greater resistance of aluminum.

Proximity Effect. *Proximity effect* in a bus is distortion of the current distribution caused by induction between the leaving and returning conductors. This distortion causes a concentration of current in the parts of the buses nearest together, thus increasing their effective resistance. The proximity effect must be taken into account for buses carrying alternating current. The effect is less on three-phase buses than on single-phase buses.

Tubular Bus. Tubular conductors used on alternating current have a better current distribution than any other shape of conductor of similar cross-sectional area, but they also have a relatively small surface area for dissipating heat losses. These two factors must be balanced properly in the design of a tubular bus.

TABLE 17-2 Current Ratings for Bare Copper Tubular Bus, Outdoors
(40°C ambient temperature, 98% conductivity copper, frequency 60 Hz, wind velocity 2 ft/s at 90° angle)

Nominal size	Outside diameter, in	Inside diameter, in	Current ratings, A		
			30°C rise	40°C rise	50°C rise
			Standard pipe sizes		
1/2	0.840	0.625	545	615	675
3/4	1.050	0.822	675	765	850
1	1.315	1.062	850	975	1080
1 1/4	1.660	1.368	1120	1275	1415
1 1/2	1.900	1.600	1270	1445	1600
2	2.375	2.062	1570	1780	1980
2 1/2	2.875	2.500	1990	2275	2525
3	3.500	3.062	2540	2870	3225
3 1/2	4.000	3.500	3020	3465	3860
4	4.500	4.000	3365	3810	4305
			Extra-heavy pipe sizes		
1/2	0.840	0.542	615	705	775
3/4	1.050	0.736	760	875	970
1	1.315	0.951	1000	1140	1255
1 1/4	1.660	1.272	1255	1445	1600
1 1/2	1.900	1.494	1445	1650	1830
2	2.375	1.933	1830	2080	2325
2 1/2	2.875	2.315	2365	2720	3020
3	3.500	2.892	2970	3365	3710
3 1/2	4.000	3.358	3380	3860	4255
4	4.500	3.818	3840	4350	4850

Note: 1 in = 25.4 mm; 1 ft/s = 0.3048 m/s.
Source: From Anderson Electric Technical Data, Table 13.

Tubing provides a relatively large cross-sectional area in minimum space and has the maximum structural strength for equivalent cross-sectional area, permitting longer distances between supports. In outdoor substations, spans of up to 40 and 50 ft with 6-in-diameter copper or aluminum tubes are considered practicable. The use of long spans reduces the number of insulator posts to a minimum. Current-carrying capacities of copper and aluminum tubular buses of different dimensions are shown in Tables 17-2 and 17-3.

Thermal Expansion. Thermal expansion and contraction of bus conductors is an important factor in bus design, particularly where high-current buses or buses of long lengths are involved. An aluminum bus will expand 0.0105 in/ft of length for a temperature rise of 38°C (100°F). In order to protect insulator supports, disconnecting switches, and equipment terminals from the stresses caused by this expansion, provisions should be made by means of expansion joints and bus-support clamps, which permit the tubing to slide.

Bus Vibration. Long tubular-bus spans have experienced vibration caused by wind blowing across the bus. Over time, this vibration can damage the bus and the equipment connected to the bus. The vibration can be eliminated or reduced by inserting a length of cable inside the tubular bus.

Bus Spacing. The spacing of buses in substations is largely a matter of design experience. However, in an attempt to arrive at some standardization of practices, minimum electrical clearances for standard basic insulation levels were established and published by the AIEE Committee on

TABLE 17-3 Current Ratings for Bare Aluminum Tubular Bus, Outdoors
(Ratings based on 30°C over 40°C ambient, frequency 60 Hz, wind velocity 2 ft/s crosswind)

Nominal size	Outside diameter, in	Inside diameter, in	Current ratings, A	
			6063-T6*	6061-T6†
ASA Schedule 40 (standard pipe size)				
1/2	0.840	0.622	405	355
3/4	1.050	0.824	495	440
1	1.315	1.049	650	575
1 1/4	1.660	1.380	810	720
1 1/2	1.900	1.610	925	820
2	2.375	2.067	1150	1020
2 1/2	2.875	2.469	1550	1370
3	3.500	3.068	1890	1670
3 1/2	4.000	3.548	2170	1920
4	4.500	4.026	2460	2180
5	5.563	5.047	3080	2730
ASA Schedule 80 (extra-heavy pipe size)				
1/2	0.840	0.546	455	400
3/4	1.050	0.742	565	500
1	1.315	0.957	740	655
1 1/4	1.660	1.278	930	825
1 1/2	1.900	1.500	1070	945
2	2.375	1.939	1350	1200
2 1/2	2.875	2.323	1780	1580
3	3.500	2.900	2190	1940
3 1/2	4.000	3.364	2530	2240
4	4.500	3.826	2880	2560
5	5.563	4.813	3640	3230

*6063-T6 = 53% IACS typical.
†6061-T6 = 40% IACS typical.
Note: 1 in = 25.4 mm.
Source: Data from Aluminum Company of America.

Substations. The data are summarized in AIEE Paper 54-80, which appeared in *Transactions* (June 1954, p. 636). This guide, shown in Table 17-4, provides minimum clearance recommendations for electric transmission systems designed for impulse-withstand levels up to and including 1175 kV BIL.

Ongoing studies attempt to extend the clearance recommendations to include the EHV range. The data published in 1954 are satisfactory to withstand anticipated switching-surge requirements of electric systems rated 161 kV and below. For systems rated 230 kV and above, more accurate determination of the switching-surge characteristics of insulation systems was required before final clearance recommendations could be made.

17.1.7 Clearance Requirements

In 1972, the Substations Committee of the IEEE published Trans. Paper T72 131-6, which established recommendations for minimum line-to-ground electrical clearances for EHV substations based on switching-surge requirements. The recommendations are based on a study of actual test data of the switching-surge strength characteristics of air gaps with various electrode configurations as reported by many investigators. The results are shown in Table 17-5 and include minimum line-to-ground clearances for EHV system voltage ratings of 345, 500, and 765 kV. The clearances given in Table 17-4 are considered adequate for both line-to-ground and phase-to-phase values for the voltage

TABLE 17-4 Minimum Electrical Clearances for Standard BIL Outdoor Alternating Current

kV class[a]	BIL level, kV withstand[b]	Minimum clearance to ground for rigid parts, in[c]	Minimum clearance between phases (or live parts) for rigid parts, in, metal to metal[d]	Minimum clearance between overhead conductors and grade for personnel safety inside substation, ft[e]	Minimum clearance between wires and roadways, inside substation enclosure, ft
7.5	95	6	7	8	20
15	110	7	12	9	20
23	150	10	15	10	22
34.5	200	13	18	10	22
46	250	17	21	10	22
69	350	25	31	11	23
115	550	42	53	12	25
138	650	50	62	13	25
161	750	58	72	14	26
230	825	65	80	15	27
230	900	71	89	15	27
	1050	83	105	16	28
	1175	94	113	17	29

[a] Coordinate kV class and BIL when choosing minimum clearances.
[b] The values above are recommended minimums but may be decreased in line with good practice, depending on local conditions, procedures, etc.
[c] The values above apply to 3300 ft above sea level. Above this elevation, the values should be increased according to IEEE Standard C37.30-1992.
[d] These recommended minimum clearances are for rigid conductors. Any structural tolerances, or allowances for conductor movement, or possible reduction in spacing by foreign objects should be added to the minimum values.
[e] These minimum clearances are intended as a guide for the installation of equipment in the field only, and not for the design of electric devices or apparatus, such as circuit breakers and transformers. 1 in = 25.4 mm; 1 ft = 0.3048 m.

classes up through 230 kV nominal system voltage where air-gap distances are dictated by impulse (BIL) withstand characteristics. The National Electric Safety Code, IEEE Standard C2-2002, also includes clearance requirements to the substation fence (Fig. 17-11).

The Substations Committee of the IEEE has an ongoing effort to review phase-to-phase air clearances and is currently balloting IEEE Standard P1427, Guide for Recommended Electrical Clearances and Insulation Levels in Air Insulated Power Substations.

Considerable information has been published by CIGRE relative to establishing phase-to-phase air clearances in EHV substations as required by switching surges. The CIGRE method is based on nearly simultaneous and equal opposite-polarity surge overvoltages in adjacent phases. The phase-to-ground surge overvoltage is multiplied by a factor of up to 1.8 (the theoretical maximum phase-to-phase voltage would be twice the phase-to-ground surge overvoltage). The estimated value of phase-to-phase overvoltage is then compared with obtained clearances. Refer to an article in CIGRE, *Electra*, no. 29, 1973, "Phase-to-Ground and Phase-to-Phase Air Clearances in Substations," by L. Paris and A. Taschini.

Suggested values of phase-to-phase clearances for EHV substations based on the CIGRE method are shown in Table 17-6. The table was formulated by choosing various phase-to-ground transient voltage values such as are used in Table 17-5. These values of phase-to-ground overvoltage were multiplied by a factor of 1.8 to arrive at a value of estimated phase-to-phase transient overvoltages. An equivalent phase-to-phase critical flashover value of voltage is next assumed by multiplying the switching-surge phase-to-phase voltage by 1.21. Finally, this value is compared with data in the CIGRE article prepared by Paris and Taschini to arrive at air-clearance values based on switching-surge impulse voltages.

EHV substation bus phase spacing is normally based on the clearance required for switching-surge impulse values plus an allowance for energized equipment projections and corona rings. This total distance may be further increased to facilitate substation maintenance.

TABLE 17-5 Minimum Electrical Clearances for EHV Substations Based on Switching Surge and Lightning Impulse Requirements
(Line to ground)

System voltage, kV		Transient voltage		SS clearances, in		BIL clearances, in	
Nom.	Max.	PU SS	Withstand SS crest, kV	Equivalent SS CFO, kV	Line to ground	Withstand BIL, kV	Line to ground
345	362	2.2	650	785	84	1050	84
		2.3	680	821	90		
		2.4	709	857	96		
		2.5	739	893	104	1300	104
		2.6	768	928	111		
		2.7	798	964	118		
		2.8	828	1000	125		
		2.9	857	1035	133		
		3.0	887	1071	140		
500	550	1.8	808	976	124	1550	124
		1.9	853	1031	132		
		2.0	898	1085	144	1800	144
		2.1	943	1139	156		
		2.2	988	1193	168		
		2.3	1033	1248	181		
		2.4	1078	1302	194		
		2.5	1123	1356	208		
		2.6	1167	1410	222		
		2.7	1212	1464	238		
		2.8	1257	1519	251		
765	800	1.5	982	1186	166	2050	167
		1.6	1047	1265	185		
		1.7	1113	1344	205		
		1.8	1178	1423	225		
		1.9	1244	1502	246		
		2.0	1309	1581	268		
		2.1	1375	1660	291		
		2.2	1440	1739	314		
		2.3	1505	1818	339		
		2.4	1571	1897	363		
		2.5	1636	1976	389		
		2.6	1702	2055	415		

Notes:

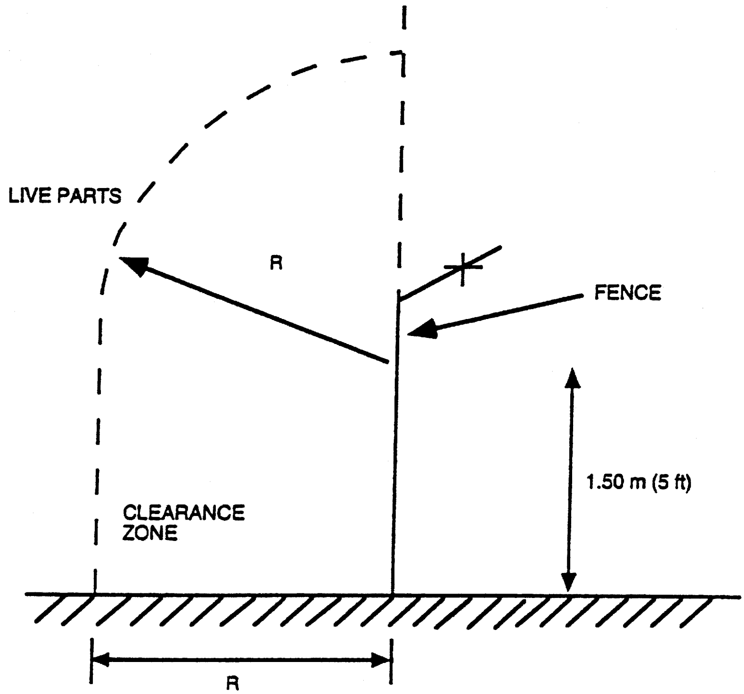
1. Minimum clearances should satisfy either maximum switching-surge or BIL duty requirement, whichever dictates the larger dimension.
2. For installations at altitudes in excess of 3300 ft elevation, it is suggested that correction factors, as provided in IEEE C37.30-1992, be applied to withstand voltages as given above.

SS: switching surge
CFO: critical flashover
1 in = 25.4 mm.

17.1.8 Mechanical and Electrical Forces

A station bus must have sufficient mechanical strength to withstand short-circuit stresses. Two factors are involved: (1) the strength of the insulators and their supporting structure and (2) the strength of the bus conductor.

A simple guide for the calculation of electromagnetic forces exerted on buses during short-circuit conditions is stated in ANSI Standard C37.32-2002, High-Voltage Switches, Bus Supports, and Accessories Schedules of Preferred Ratings, Construction Guidelines and Specifications.



Nominal voltage between phases	Typical BIL	Dimension R m	Dimension R ft
151-7200	95	3.0	10.0
13,800	110	3.1	10.1
23,000	150	3.1	10.3
34,500	200	3.2	10.6
46,000	250	3.3	10.9
69,000	350	3.5	11.6
115,000	550	4.0	13.0
138,000	650	4.2	13.7
161,000	750	4.4	14.3
230,000	825	4.5	14.9
230,000	900	4.7	15.4
345,000	1050	5.0	16.4
500,000	1175	5.3	17.3

FIGURE 17-11 Substation fence clearance requirements. *(National Electrical Safety Code, IEEE C2-2002.)*

The electromagnetic force exerted between two current-carrying conductors is a function of the current, its decrement rate, the shape and arrangement of conductors, and the natural frequencies of the complete assembly, including mounting structure, insulators, and conductors. Obviously, it is not feasible to cover each and every case with one simple equation, even if some approximations are made, because of the large number of variables involved, including the wide range of constants for support structures.

The force calculated by the following equation is that produced by the maximum peak current. In most cases, the calculated force is higher than that which actually occurs, due to inertia and flexibility

TABLE 17-6 Suggested Electrical Clearances for EHV Substations Based on Switching Surge Requirements and Including U.S. Utility Practice
(Phase to phase)

System voltage, kV		Transient voltage		SS clearances, in*		Present practice
Nom.	Max.	SS L-G PU	Withstand SS crest, kV	Equivalent L-L SS CFO, kV	Rod to rod withstand*	U.S. utility phase spacing, ft
345	362	2.2	650	1405	103	15 to 18
		2.6	768	1660	128	
		3.0	887	1915	159	
500	550	1.8	808	1745	138	20 to 35
		2.2	988	2135	190	
		2.5	1123	2425	239	
		2.8	1257	2715	294	
765	800	1.5	982	2120	189	45 to 50
		1.8	1178	2545	261	
		2.1	1375	2970	356	
		2.4	1571	3395	480	

Note: 1 in = 25.4 mm; 1 ft = 0.3048 m.

*The values of L-L switching-surge clearances are based on the use of SS L-G crest voltages multiplied by 1.8. This value of L-L SS voltage is then multiplied by 1.21 to indicate an SS CFO value of voltage used to determine the clearances. For a description of method used, refer to CIGRE report by L. Paris and A. Taschini, Phase-to-Ground and Phase-to-Phase Air Clearances in Substations, CIGRE, *Electra*, no. 29, 1973, pp. 29–44. L-G: line-to-ground; L-L: line-to-line; SS: switching surge; CFO: critical flashover.

of the systems, and this fact tends to compensate for the neglect of resonant forces. The equation, therefore, is sufficiently accurate for usual practice conditions.

$$F = M\frac{5.5 \times I^2}{S \times 10^7} \tag{17-1}$$

where F = pounds per foot of conductor
M = multiplying factor
I = short-circuit current, A (defined in Table 17-7)
S = spacing between centerlines of conductors, in

After determining the value of I, select the corresponding M factor from Table 17-7.

Structures with long spans held in tension by strain insulators cannot be calculated for stresses by the preceding procedure, but approximate estimates can be made by following the procedure generally used for calculating mechanical stresses in transmission-line conductors.

TABLE 17-7 Multiplying Factor (M) for Calculation of Electromagnetic Forces

Circuit	Amperes (I) expressed as	Multiplying factor (M)
dc	Max. peak	1.0
ac, 3-phase	Max. peak	0.866
ac, 3-phase	rms asymmetrical	$(0.866 \times 1.63^2) = 2.3$
ac, 3-phase	rms symmetrical	$(0.866 \times 2.82^2) = 6.9$
1 phase of 3 phase or 1 phase	Max peak	1.0
1 phase of 3 phase or 1 phase	rms asymmetrical	$(1.63^2) = 2.66$
1 phase of 3 phase or 1 phase	rms symmetrical	$(2.82^2) = 8.0$

The total stress in an outdoor bus is the resultant of the stresses due to the short-circuit load together with the dead, ice, and wind loads.

1. *Buses up to 161 kV.* The distance between phases and the character of the bus supports and their spacing are such that wind loading usually may be neglected. Ice load of $^1/_2$ in is usually considered.
2. *Buses for 230 kV and higher voltages.* The spacing between phases is usually so large that the mechanical effects of short-circuit currents may not be the determining factor, and such buses, when designed properly for the mechanical loads only, may be found to also satisfy the electrical short-circuit current requirements. However, short-circuit duties on modern systems continue to rise, and the electrical forces should be checked by Eq. (17-1).

Deflections and stresses on aluminum buses can be determined by referring to Tables 17-8 and 17-9. All loads are assumed to be uniformly distributed. Loading includes the dead load of the bus and, in addition, includes ice loadings of $^1/_2$- and 1-in coating on the bus. Wind loads are assumed to be 8 lb/ft^2 of the projected area of tubing including $^1/_2$ in of ice. Large deflections should be avoided even if the maximum bending stress is found to be within safe limits. It is generally satisfactory, in approximation of bus diameter, to allow 1 in of bus outside diameter for every 10 ft of bus span. Refer to the foot notes below Tables 17-8 and 17-9 for the method of support and number of spans.

Stresses on disconnecting switches under short-circuit conditions may be sufficient to open them, with disastrous results; therefore, modern switch designs embody locks, or overtoggle mechanisms, to prevent this from occurring. The force on the switchblade varies as the square of the current. This force will be increased if the return circuit passes behind the switch and will vary inversely with the distance from the center of the switchblade to the center of the return conductor.

Bus supports are designed for definite cantilever strength, expressed in inch pounds and measured at the cap supporting the conductor clamp. Ample margin of safety with regard to insulation and structural strength should be provided, manufacturers' data should be checked carefully, and units should be so selected that allowable values for the particular units are not exceeded. Good practice recommends that the working load must not exceed 40% of the published rating, and short-circuit loads must not exceed the insulator published rating. These loads should include forces for ultimate short-circuit growth and worst mechanical loading.

17.1.9 Overvoltage and Overcurrent Protection

Protective Relaying. A substation can employ many relaying systems to protect the equipment associated with the station; the most important of these are

1. Transmission and distribution lines emanating from the station
2. Step-up and step-down transformers
3. Station buses
4. Breakers
5. Shunt and series reactors
6. Shunt and series capacitors

Substations serving bulk transmission system circuits must provide a high order of reliability and security in order to provide continuity of service to the system. More and more emphasis is being placed on very sophisticated relaying systems which must function reliably and at high speeds to clear line and station faults while minimizing false tripping.

Most EHV and UHV systems now use two sets of protective relays for lines, buses, and transformers. Many utilities use one set of electromechanical relays for transmission-line protection, with a completely separate, redundant set of solid-state relays to provide a second protective relaying package or two completely separate redundant sets of solid-state relays. The use of two separate sets of relays, operating from separate potential and current transformers and from separate station batteries, allows for the testing of relays without the necessity of removing the protected line or bus from service. For more difficult relaying applications, such as EHV lines using series

TABLE 17-8 Aluminum Round Tubular Bus bar Deflections and Stresses
(Standard iron pipe sizes)

IPS size, in	Loading	Span, ft 20 Deflection, in	20 Stress, lb/in²	25 Deflection, in	25 Stress, lb/in²	30 Deflection, in	30 Stress, lb/in²	35 Deflection, in	35 Stress, lb/in²	40 Deflection, in	40 Stress, lb/in²	45 Deflection, in	45 Stress, lb/in²	50 Deflection, in	50 Stress, lb/in²
1 1/4	Bare	1.45	2010	3.54	3135										
	1/2″ ice	3.94	5445	9.61	8510										
	1/2″ ice + 8 lb wind	5.12	7090	12.51	11075										
	1″ ice	7.57	10470	18.48	16360										
1 1/2	Bare	1.09	1725	2.66	2700										
	1/2″ ice	2.83	4475	6.90	6990										
	1/2″ ice + 8 lb wind	3.61	5715	8.81	8930										
	1″ ice	5.28	8365	12.90	13070										
2	Bare	0.68	1350	1.67	2110	3.45	3040								
	1/2″ ice	1.65	3265	4.03	5100	8.35	7345								
	1/2″ ice + 8 lb wind	2.05	4055	5.00	6340	10.38	9125								
	1″ ice	2.95	5845	7.21	9135	14.95	13150								
2 1/2	Bare	0.47	1130	1.15	1765	2.38	2540	4.42	3455						
	1/2″ ice	0.96	2310	2.36	3610	4.89	5200	9.05	7080						
	1/2″ ice + 8 lb wind	1.14	2730	2.78	4270	5.77	6150	10.70	8370						
	1″ ice	1.61	3845	3.92	6010	8.13	8655	15.06	11780						
3	Bare	0.31	910	0.76	1425	1.58	2050	2.93	2790	5.00	3640				
	1/2″ ice	0.61	1775	1.49	2775	3.08	3995	5.71	5440	9.74	7105				
	1/2″ ice + 8 lb wind	0.71	2060	1.72	3220	3.58	4635	6.62	6310	11.30	8240				
	1″ ice	0.98	2860	2.39	4465	4.96	6430	9.19	8755	15.68	11435				
3 1/2	Bare	0.24	790	0.58	1230	1.20	1775	2.22	2415	3.79	3155	6.06	3995		
	1/2″ ice	0.45	1490	1.09	2330	2.26	3355	4.19	4565	7.15	5960	11.46	7545		
	1/2″ ice + 8 lb wind	0.51	1710	1.25	2670	2.59	3845	4.81	5230	8.20	6835	13.14	8650		
	1″ ice	0.70	2350	1.72	3670	3.57	5280	6.61	7190	11.27	9390	18.05	11885		

4	Bare	0.19	695	0.45	1090	0.94	1565	1.74	2130	2.97	2785	4.76	3525	7.25	4350
	1/2″ ice	0.34	1275	0.83	1995	1.72	2870	3.19	3910	5.45	5105	8.72	6465	13.30	7980
	1/2″ ice + 8 lb wind	0.39	1450	0.94	2265	1.96	3260	3.62	4435	6.18	5795	9.90	7335	15.09	9055
	1″ ice	0.53	1975	1.28	3085	2.66	4440	4.93	6045	8.42	7895	13.49	9990	20.55	12330
4 1/2	Bare	0.15	620	0.36	970	0.76	1400	1.40	1905	2.39	2490	3.83	3150	5.83	3890
	1/2″ ice	0.27	1115	0.65	1740	1.35	2505	2.51	3410	4.28	4455	6.85	5640	10.44	6960
	1/2″ ice + 8 lb wind	0.30	1255	0.73	1960	1.52	2820	2.82	3840	4.81	5015	7.71	6345	11.75	7835
	1″ ice	0.41	1695	0.99	2650	2.06	3810	3.81	5190	6.51	6780	10.42	8580	15.89	10590
5	Bare	0.12	555	0.29	870	0.61	1250	1.12	1705	1.92	2225	3.07	2815	4.69	3475
	1/2″ ice	0.21	970	0.51	1520	1.06	2185	1.96	2975	3.35	3885	5.37	4920	8.18	6070
	1/2″ ice + 8 lb wind	0.23	1085	0.57	1695	1.18	2440	2.19	3320	3.74	4335	5.99	5490	9.13	6775
	1″ ice	0.31	1455	0.77	2275	1.59	3275	2.94	4455	5.02	5820	8.04	7365	12.26	9095
6	Bare	0.08	465	0.20	725	0.42	1040	0.79	1420	1.34	1850	2.15	2345	3.28	2895
	1/2″ ice	0.14	775	0.34	1210	0.71	1745	1.32	2375	2.25	3105	3.60	3930	5.49	4850
	1/2″ ice + 8 lb wind	0.15	855	0.38	1335	0.78	1925	1.45	2615	2.48	3420	3.97	4325	6.05	5340
	1″ ice	0.21	1135	0.50	1770	1.04	2550	1.92	3470	3.28	4530	5.26	5735	8.02	7080

Note: The tabulated deflections are for single-span, simply supported buses. Deflections for fixed-end buses are one-fifth of the values given above, and the deflections for continuous buses for the center spans are also one-fifth of the values above. The deflections for the end spans are two-fifths of the values given. The stresses given in the above table are the stresses in the outer fibers as calculated for simply supported beams with a uniformly distributed load. 1 in = 25.4 mm; 1 ft = 0.3048 m; 1 lb = 0.4536 kg; 1 lb/in^2 = 6.895 kPa.

Source: From *Kaiser Aluminum Electrical Conductor Technical Manual*.

TABLE 17-9 Aluminum Round Tubular Bus bar Deflections and Stresses
(Extra-heavy pipe sizes)

IPS size, in	Loading	Span, ft: 20 Deflection, in	20 Stress, lb/in^2	25 Deflection, in	25 Stress, lb/in^2	30 Deflection, in	30 Stress, lb/in^2	35 Deflection, in	35 Stress, lb/in^2	40 Deflection, in	40 Stress, lb/in^2	45 Deflection, in	45 Stress, lb/in^2	50 Deflection, in	50 Stress, lb/in^2
1¼	Bare	1.54	2130												
	½″ ice	3.54	4900												
	½″ ice + 8 lb wind	4.42	6110												
	1″ ice	6.47	8945												
1½	Bare	1.15	1825	2.82	2855										
	½″ ice	2.53	4005	6.17	6255										
	½″ ice + 8 lb wind	3.09	4895	7.55	7645										
	1″ ice	4.47	7085	10.92	11070										
2	Bare	0.72	1425	1.76	2225										
	½″ ice	1.46	2890	3.57	4520										
	½″ ice + 8 lb wind	1.73	3430	4.23	5360										
	1″ ice	2.46	4870	6.01	7610										
2½	Bare	0.49	1185	1.21	1850	2.50	2665								
	½″ ice	0.89	2125	2.17	3320	4.49	4780								
	½″ ice + 8 lb wind	1.01	2420	2.46	3780	5.11	5440								
	1″ ice	1.40	3345	3.41	5225	7.07	7525								
3	Bare	0.33	955	0.80	1495	1.66	2150	3.07	2925						
	½″ ice	0.56	1625	1.36	2540	2.82	3660	5.23	4980						
	½″ ice + 8 lb wind	0.62	1815	1.52	2840	3.15	4085	5.84	5560						
	½″ ice	0.85	2465	2.06	3850	4.28	5545	7.93	7550						
3½	Bare	0.25	825	0.60	1290	1.25	1860	2.32	2530	3.96	3305				
	½″ ice	0.41	1360	1.00	2125	2.07	3060	3.83	4165	6.53	5440				
	½″ ice + 8 lb wind	0.45	1500	1.10	2345	2.28	3380	4.23	4600	7.21	6010				
	1″ ice	0.60	2015	1.48	3145	3.06	4530	5.67	6170	9.67	8055				

4	Bare	0.19	725	0.47	1135	0.98	1635	1.82	2230	3.10	2910	4.97	3680	7.58	4545
	1/2″ ice	0.31	1165	0.76	1820	1.57	2620	2.91	3565	4.97	4655	7.96	5895	12.13	7275
	1/2″ ice + 8 lb wind	0.34	1270	0.83	1990	1.72	2865	3.18	3900	5.43	5095	8.70	6445	13.26	7955
	1″ ice	0.45	1690	1.10	2640	2.28	3800	4.22	5170	7.20	6755	11.54	8550	17.59	10555
4 1/2	Bare	0.16	650	0.38	1015	0.79	1460	1.46	1990	2.49	2600	4.00	3290	6.09	4060
	1/2″ ice	0.24	1015	0.59	1585	1.23	2285	2.28	3110	3.90	4060	6.24	5135	9.51	6340
	1/2″ ice + 8 lb wind	0.26	1100	0.65	1720	1.34	2475	2.48	3370	4.23	4405	6.77	5575	10.32	6880
	1″ ice	0.35	1445	0.85	2260	1.76	3255	3.26	4430	5.55	5785	8.90	7320	13.56	9040
5	Bare	0.13	580	0.31	905	0.63	1305	1.17	1775	2.00	2320	3.20	2935	4.88	3625
	1/2″ ice	0.19	885	0.47	1380	0.97	1990	1.79	2710	3.05	3535	4.89	4475	7.45	5525
	1/2″ ice + 8 lb wind	0.21	950	0.50	1490	1.04	2140	1.93	2915	3.29	3810	5.26	4820	8.02	5950
	1″ ice	0.27	1240	0.65	1935	1.35	2785	2.51	3795	4.28	4955	6.85	6270	10.44	7745
6	Bare	0.09	485	0.21	755	0.44	1090	0.82	1485	1.40	1940	2.25	2455	3.43	3030
	1/2″ ice	0.13	700	0.31	1095	0.64	1580	1.19	2150	2.04	2810	3.26	3555	4.97	4390
	1/2″ ice + 8 lb wind	0.13	745	0.33	1165	0.68	1675	1.27	2280	2.16	2980	3.46	3775	5.27	4660
	1″ ice	0.17	950	0.42	1485	0.87	2140	1.61	2910	2.75	3800	4.41	4810	6.72	5940

Note: The tabulated deflections are for single-span, simply supported buses. Deflections for fixed-end buses are one-fifth of the values given above, and the deflections for continuous buses for the center spans are also one-fifth of the values above. The deflections for the end spans are two-fifths of the values given. The stresses given in the above table are the stresses in the outer fibers as calculated for simply supported beams with a uniformly distributed load. 1 in = 25.4 mm; 1 ft = 0.3048 m; 1 lb = 0.4536 kg; 1 lb/in^2 = 6.895 kPa.

Source: From *Kaiser Aluminum Electrical Conductor Technical Manual*.

capacitors in the line, some companies always use two sets of solid-state relays to provide the protection systems.

Transmission-line relay terminals are located at the substation and employ many different types of relaying schemes that include the following:

1. Pilot wire
2. Direct underreaching
3. Permissive underreaching
4. Permissive overreaching
5. Directional comparison
6. Phase comparison

Pilot-Wire Relaying. Pilot-wire relaying is an adaptation of the principle of differential relaying to line protection and functions to provide high-speed clearing of the line for faults anywhere on the line. Pilots include wire pilot (using a two-wire pair between the ends of the line), carrier-current pilots, microwave pilots, fiber-optics pilots, and the use of audio-tone equipment over wire, carrier, fiber-optics, or microwave. The transmission lines may have two or more terminals each with circuit breakers for disconnecting the line from the rest of the power system. All the relaying systems described can be used on two-terminal or multiterminal lines. The relaying systems program the automatic operation of the circuit breakers during power-system faults.

Direct Underreaching Fault Relays. These relays (Fig. 17-12) at each terminal of the protected line sense fault power flow into the line. Their zones of operation must overlap but not overreach any remote terminals. The operation of the relays at any terminal initiates both the opening of the local breaker and the transmission of a continuous remote tripping signal to effect instantaneous operation of all remote breakers. For example, in Fig. 17-12, for a line fault near bus *A*, the fault relays at *A* open (trip) breaker *A* directly and send a transfer trip signal to *B*. The reception of this trip signal at *B* trips breaker *B*.

Permissive Underreaching Relays. The operation and equipment for this system are the same as those of the direct underreaching system, with the addition of fault-detector units at each terminal. The fault detectors must overreach all remote terminals. They are used to provide added security by supervising remote tripping. Thus, the fault relays operate as shown in Fig. 17-12 and the fault detectors as shown in Fig. 17-13. As an example, for a fault near *A* in Fig. 17-12, the fault relays at *A* trip breaker *A* directly and send a transfer trip signal to *B*. The reception of the trip signal plus the operation of the fault detector relays at *B* (Fig. 17-13) trip breaker *B*.

Permissive Overreaching Relays. Fault relays at each terminal of the protected line sense fault power flow into the line, with their zones of operation overreaching all remote terminals. Both the

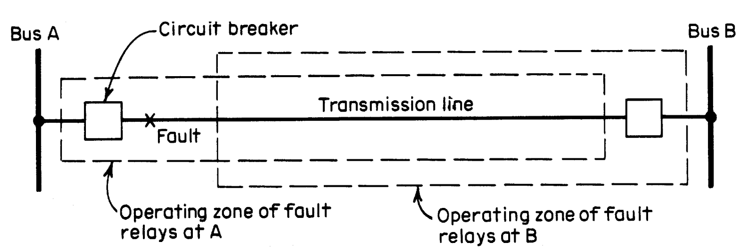


FIGURE 17-12 Fault-relay operating zones for the underreaching transfer trip transmission-line pilot relaying system.

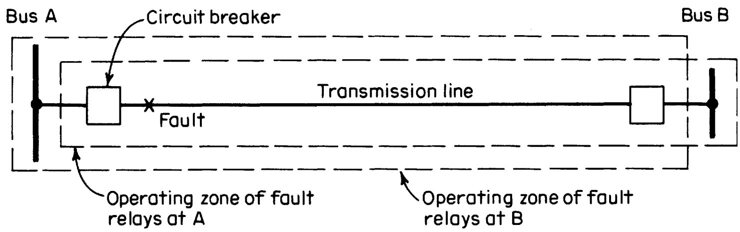


FIGURE 17-13 Fault-relay operating zones for the overreaching transmission-line pilot relaying system.

operation of the local fault relays and a transfer trip signal from all the remote terminals are required to trip any breaker. Thus, in the example of Fig. 17-13 for the line fault near *A*, fault relays at *A* operate and transmit a trip signal to *B*. Similarly, the relays at *B* operate and transmit a trip signal to *A*. Breaker *A* is tripped by the operation of the fault relay *A* plus the remote trip signal from *B*. Likewise, breaker *B* is tripped by the operation of fault relay *B* plus the remote trip signal from *A*.

Directional-Comparison Relays. The channel signal in these systems (Fig. 17-14) is used to block tripping in contrast to its use to initiate tripping in the preceding three systems. Fault relays at each terminal of the protected line section sense fault power flow into the line. Their zones of operation must overreach all remote terminals. Additional fault-detecting units are required at each terminal to initiate the channel-blocking signal. Their operating zones must extend further or be set more sensitively than the fault relays at the far terminals. For example, in Fig. 17-13 the blocking zone at *B* must extend further behind breaker *B* (to the right) than the operating zone of the fault relays at *A*. Correspondingly, the blocking zone at *A* must extend further out into the system (to the left) than the operating zone of the fault relays at *B*.

For an internal fault on line *AB*, no channel signal is transmitted (or if transmitted, it is cut off by the fault relays) from any terminal. In this absence of any channel signal, fault relays at *A* instantly trip breaker *A*, and fault relays at *B* instantly trip breaker *B*. For the external fault to the right of *B* as shown in Fig. 17-13, the blocking zone relays at *B* transmit a blocking channel signal to prevent the fault relays at *A* from tripping breaker *A*. Breaker *B* is not tripped because the *B* operating zone does not see this fault.

Phase-Comparison Relays. The three line currents at each end of the protected line are converted into a proportional single-phase voltage. The phase angles of the voltages are compared by permitting

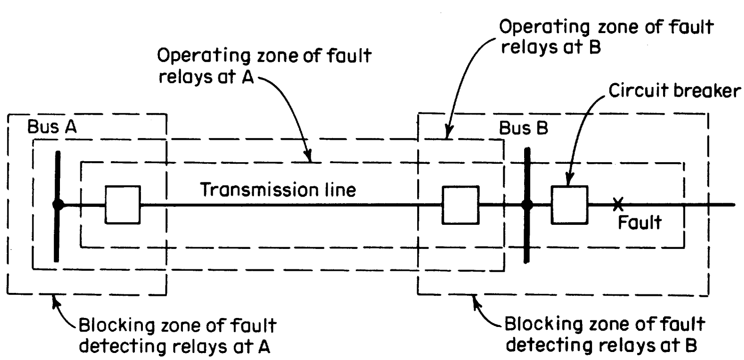


FIGURE 17-14 Fault and blocking relay operating zones for the directional-comparison transmission-line pilot relay system.

the positive half-cycle of the voltage to transmit a half-wave signal block over the pilot channel. For external faults, these blocks are out of phase so that alternately the local and then the remote signal provide essentially a continuous signal to block or prevent tripping. For internal faults, the local and remote signals are essentially in phase so that approximately a half-cycle of no channel signal exists. This is used to permit the fault relays at each terminal to trip their respective breakers.

Station Bus Protection. Station bus protection deserves very careful attention because bus failures are, as a rule, the most serious that can occur to an electrical system. Unless properly isolated, a bus fault could result in complete shutdown of a station. Many methods are employed to protect the station buses. Among them are the use of overcurrent relays, backup protection by relays of adjacent protective zones, directional-comparison schemes, and so forth. By far, the most effective and preferred method used to protect buses consists of percentage differential relaying, using either current or voltage differential relays. Differential relaying is preferred because it is fast, selective, and sensitive.

The relays are available in either electromechanical or solid-state form, with the solid-state units featuring somewhat higher speeds and sensitivity than are available in the electromechanical models. Operating times of 5 to 8 ms can be achieved with solid-state bus differential relays.

Because of the high magnitude of currents encountered during bus faults, current transformers may saturate and thus cause false tripping during external faults. The possibility of ac and dc saturation during faults makes it mandatory that current transformers used for bus differential protection be accurate and of the best quality possible. Also, current transformers should be matched to provide similar ratios and characteristics.

Some bus differential relays developed in solid-state form in Europe have been designed to function correctly even when using current transformers of inferior quality and different ratios. However, it is considered good practice to provide the best possible current transformers for use in bus differential relay applications. For a sensitive bus differential scheme using current percentage differential relays, refer to Fig. 17-15. For a percentage differential scheme using high-impedance-voltage differential relays, refer to Fig. 17-16.

Because the effective resistance of the voltage relay coil circuit is so high, of the order of 3000 Ω, a voltage-limiting element must be connected in parallel with the rectifier branch in order to prevent the CT secondary voltage from being excessive. The overcurrent relay in series with the voltage limiter provides high-speed operation for bus faults of high currents. All current-transformer leads are paralleled at a junction point in the substation near the circuit breakers, and only one set of leads is required to be run into the control house where the relay is normally located.

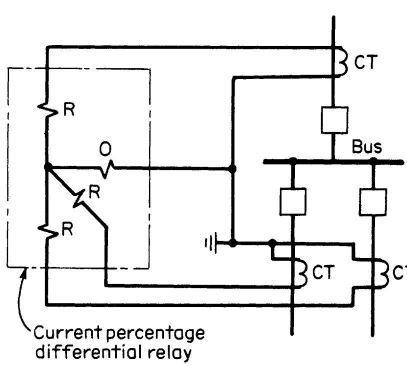


FIGURE 17-15 Bus differential protection using current percentage differential relays (CT, current transformer; O, operating coil; R, restraining coil).

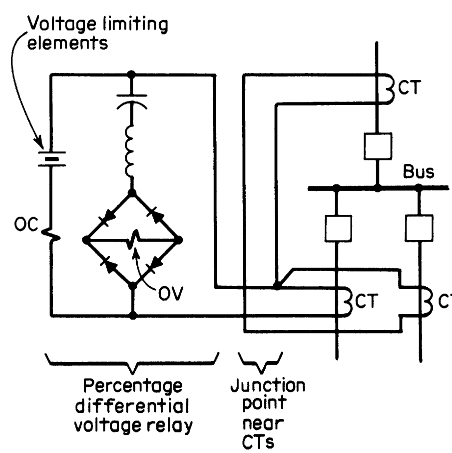


FIGURE 17-16 Bus differential protection using voltage differential relays; (OC, high set overcurrent relays; OV, voltage element; CT, current transformer).

Transformer Protection. Transformers may be subjected to short circuits between phase and ground, open circuits, turn-to-turn short circuits, and overheating. Interphase short circuits are rare and seldom develop as such initially, since the phase windings are usually well separated in a three-phase transformer. Faults usually begin as turn-to-turn failures and frequently develop into faults involving ground.

It is highly desirable to isolate transformers with faulty windings as quickly as possible to reduce the possibility of oil fires, with the attendant resulting cost for replacement. Differential protection is the preferred type of transformer protection due to its simplicity, sensitivity, selectivity, and speed of operation. If the current-transformer ratios are not perfectly matched, taking into account the voltage ratios of the transformer, autotransformers or auxiliary current transformers are required in the current-transformer secondary circuits to match the units properly so that no appreciable current will flow in the relay operating coil, except for internal fault conditions.

In applying differential protection to transformers, somewhat less sensitivity in the relays is usually required, as compared with generator relays, since they must remain nonoperative for the maximum transformer tap changes that might be used. It is also necessary to take into account the transformer exciting inrush current that may flow in only one circuit when the transformer is energized by closing one of its circuit breakers. As a rule, incorrect relay operation can be avoided by imposing a slight time delay for this condition.

Voltage-load tap-changing (LTC) transformers may be protected by differential relays. The same principles of applying differential protection to other transformers hold here as well. It is important that the differential relay be selected carefully so that the unbalance in the current-transformer secondary circuits will not in any case be sufficient to operate the relay under normal conditions. It is suggested that the current transformers be matched at the midpoint of the tap-changing range. The current-transformer error will then be a minimum for the maximum tap position in either direction.

Current-transformer and relay connections for various types of differential protection are indicated (1) in Fig. 17-17 for a Y-delta transformer and (2) in Fig. 17-18 for a three-winding Y-delta-Y transformer. Two rules, frequently used in laying out the wiring for differential protection of transformers whose main windings are connected in Y and delta, are

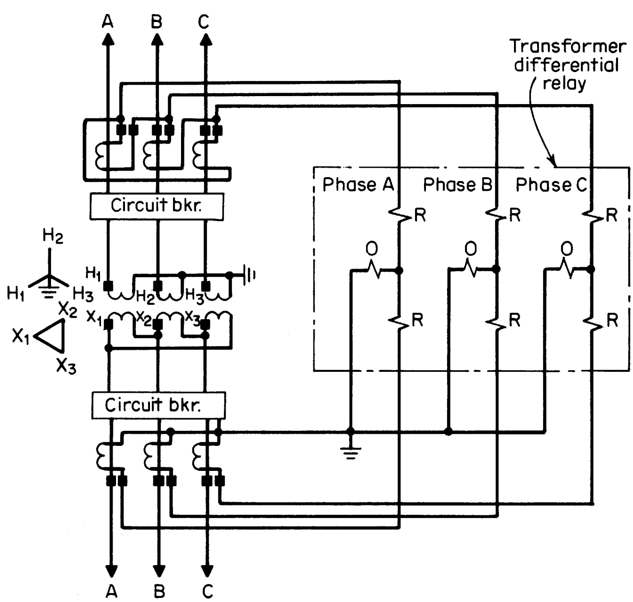


FIGURE 17-17 Transformer differential protection for a Y-Δ transformer.

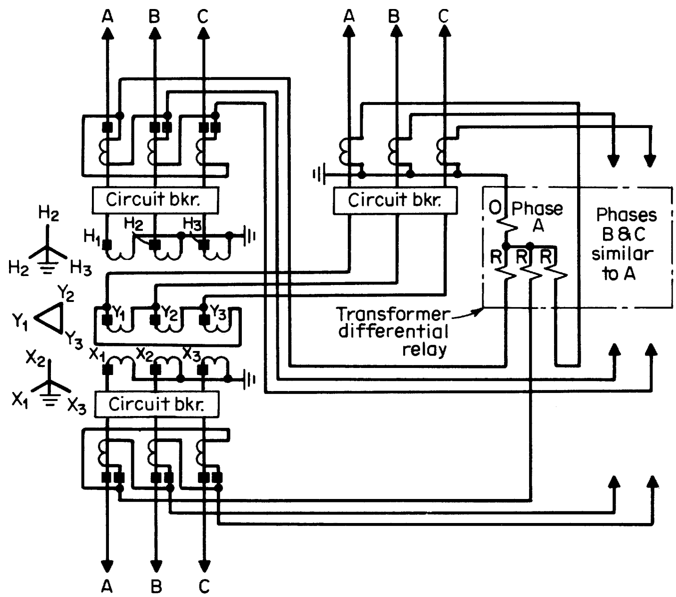


FIGURE 17-18 Transformer differential protection for a Y-Δ Y transformer.

1. The current transformers in the leads to the Y-connected winding should be connected in delta; current transformers in the leads to a delta-connected winding should be connected in Y.
2. The delta connection of the current transformers should be a replica of the delta connection of the power transformers; the Y connection of the current transformers should be a replica of the Y connection of the power transformers.

Current transformers that will give approximately 5-A secondary current at full load on the transformer should be chosen. This will not be possible in all cases, particularly for transformers having three or more windings, since the kVA ratings may vary widely and may not be proportional to the voltage ratings.

Overcurrent protection should be applied to transformers as the primary protection where a differential scheme cannot be justified or as "backup" protection if differential is used. Frequently, faster relaying may be obtained for power flow from one direction by the use of power-directional relays.

Transformer overheating protection is sometimes provided to give an indication of overtemperature, rarely to trip automatically. Overload relays of the replica type may be connected in the current-transformer circuits to detect overloading of the unit. Others operate on top-oil temperature, and still others operate on top-oil temperature supplemented with heat from an adjacent resistor connected to a current transformer in the circuit. A recently developed sensor using a glass chip sensitive to temperature changes employs fiber-optics techniques to measure winding hot-spot temperatures.

Gas- or oil-pressure relays are available for attachment to the top or side of transformer tanks to indicate winding faults, which produce gas or sudden pressure waves in the oil. Rapid collection of gas or pressure waves in the oil, due to short circuits in the winding, will produce fast operation. New, more sophisticated methods to detect incipient failures by frequent monitoring of gas samples are being developed.

Circuit-Breaker Protection. In recent years, great emphasis has been placed on the need to provide backup protection in the event of failure of a circuit breaker to clear a fault following receipt of a trip command from protective relays. For any fault, the protective relays operate to trip the necessary circuit breakers. In addition, these same protective relays, together with breaker-failure fault-detector relays, will energize a timer to start the breaker-failure backup scheme. If any breaker fails to clear the fault, the protective relays will remain picked up, permitting the timer to time-out and trip the necessary other breakers to clear the fault.

Circuit-breaker failure can be caused by loss of dc trip supply, blown trip fuses, trip-coil failure, failure of breaker trip linkages, or failure of the breaker current-interrupting mechanism. The two basic types of failures are (1) mechanical failure and (2) electrical failure of the breaker to clear the fault. Mechanical failure occurs when the breaker does not move following receipt of a trip command because of loss of dc trip supply, trip coil failure, or trip linkage failure. Electrical failure occurs when the breaker moves in an attempt to clear a fault on receipt of the trip command but fails to break the fault current because of misoperation of the current interrupter itself.

In order to clear faults for these two types of breaker failures, two different schemes of protection can be employed. The more conventional breaker-failure schemes consist of using instantaneous current-operated fault detectors which pick up to start a timer when fault relays operate. If the breaker fails to operate to clear the fault, the timer times out and trips necessary breakers to clear the fault. However, if the breaker operates correctly to clear the fault, enough time must be allowed in the timer setting to ensure reset of the fault-detector relay. Total clearing times at EHV using this scheme are quite fast and usually take 10 to 12 cycles from the time of fault until the fault is cleared.

For those faults where mechanical failure of the breakers occurs, an even faster scheme is in use. This scheme depends on a breaker auxiliary switch (normally open type 52-A contact) to initiate a fast timer. The auxiliary switch is specially located to operate from breaker trip linkages to sense actual movement of the breaker mechanism. If the breaker failure is mechanical, the breaker-failure timer is actuated through the auxiliary switch when the protective relays operate. The advantage of using the auxiliary switch is the extremely fast reset time of the breaker-failure timer that can be realized when the breaker operates correctly. Schemes in use with the fast breaker-failure circuit can attain total clearing times of 7.5 cycles when a breaker failure occurs.

Shielding and Grounding Practices for Control Cables. For several years, the increased application of solid-state devices for protective relaying and control and for electronic equipment, such as audio tones, carrier and microwave equipment, event recorders, and supervisory control equipment, in EHV substations has resulted in many equipment failures. Many of these failures have been attributed to transients or surges in the control circuits connected to the solid-state devices. Failures due to transients or surges have been experienced even with conventional electromechanical devices.

The failures being experienced are attributed to the use of EHV (345 kV and higher voltage levels) as well as the presence of unusually high short-circuit currents. One of the major sources of transient voltages is the switching of capacitances, for example, the operation of a disconnect switch which generates high-magnitude, high-frequency oscillatory surge currents. The transient magnetic fields associated with these high-frequency surge currents are both electrostatically and magnetically coupled to cables in the area. Induced voltages have been reported to be as high as 10 kV in cables without shielding, and the frequencies of these induced voltages have been reported to be as high as 3 MHz.

In order to avoid insulation breakdown at 10-kV crest and possible false operation of relays, it is important that station design includes necessary precautions to limit the undesirable surges and control circuit transients to an acceptable minimum.

In any station design there are several precautions that can be taken. All cable circuits that are used in a substation should be run radially, with each circuit separated from any other circuit and with both supply and return conductors contained within the same cable. If a conductor is routed from the control house to a point in the switchyard with the return circuits following different paths, loops may be formed that are inductive and are subject to magnetically induced voltages. However, when the two conductors involved are both affected by the same field, the voltage appearing between them at the open end should be essentially zero.

Because of ground-mat potential differences and longitudinally induced voltages in the radial circuits, proper cable shielding is necessary to maintain lowest possible voltages on the cable leads. The cables that require shielding include control, current, and potential transformer circuits. The shield should be of as low resistance as possible, and it should be connected to the ground grid at least at both ends. To reduce penetration of a magnetic flux through the nonferrous shield (lead, copper, bronze, etc.), a current must flow in the shield to produce a counterflux, which opposes the applied flux. Ground-grid conductors should be placed in parallel to and in close proximity to the shield to maintain as low a resistance between the ends as possible and also to form a small loop to reduce the reactance between ground and the shield. Without close coupling of the conductor and ground shield, the propagation time of the two paths could differ so that a voltage impulse could arrive at the receiving end with a time difference, hence causing an unwanted voltage difference.

All control, potential-transformer, and current-transformer cables should be shielded, with the shield grounded at the switchyard end and at the control-house end. In addition, each group or run of conduits and cables should be installed with a separate No. 4/0 bare stranded copper cable buried directly in the ground and grounded and bonded to the control-cable shield at each end of each cable. The bare copper cable should run as closely as possible to the cable run. The heavy cable functions to provide a low-resistance path in an attempt to prevent heavy fault currents from flowing in the shield and to reduce reactance between ground and shield.

In order to limit induced voltages, the control-cable runs should be installed, where possible, at right angles to high-voltage buses. Where it is necessary to run parallel to a high-voltage bus for any appreciable distance, the spacing between cables and high-voltage buses should be made as great as possible. Distances of at least 50 ft should be maintained.

It is further considered good practice to have both current-transformer and potential-transformer leads installed with the ground for the secondary wye neutral made at the control-house end rather than at the switchyard end. Any rise due to induced voltages will be concentrated at the switchyard and will ensure operator safety at the control switchboard in the control house.

The shield can be grounded by using a flexible tinned copper braid of from $\frac{1}{2}$ to 1 in wide. The shielded-cable outer insulation is peeled back, exposing the sheath. The 1-in braid is wrapped around the sheath and soldered carefully to it. The other end of the braid is connected to a lug, and solder should be run over the lug to the braid connection. The lug is then bolted securely to the ground bus bar. The flexible copper braid circuits should be kept as short as possible and should be run directly to the ground bus without any bends, if possible.

It should be pointed out that the shields should be grounded at multiple points rather than at a single point, because of the tendency to lose any advantage from single-point grounding at 50 kHz and above. As an example, assume that one input and one output terminal of a system are grounded, each at different points on a common ground plane. A small noise voltage will usually exist across these ground points because of currents flowing in the finitely conductive ground plane. If either the load or source ground is lifted, a ground loop is no longer formed, and coupling of unwanted signals is minimized. This is the advantage of having one physical ground.

Removal of one of the ground connections achieves a single-point ground only for dc and low-frequency signals. At higher frequencies, ground loops will be created by capacitance coupling. Frequencies below 50 kHz are considered the arbitrary crossover point for single-point grounding. At EHV, the transient voltages above 50 kHz represent the more serious problem; for this reason, all cable shields should be grounded at least at two points. It should be noted that shielding of control cables is normally provided for substations operating at voltage levels of 138 kV and above.

17.1.10 Substation Grounding

Grounding at substations is highly important. The functions of a grounding system are listed below:

1. Provide the ground connection for the grounded neutral for transformers, reactors, and capacitors
2. Provide the discharge path for lightning rods, arresters, gaps, and similar devices
3. Ensure safety to operating personnel by limiting potential differences, which can exist in a substation
4. Provide a means of discharging and deenergizing equipment in order to proceed with maintenance on the equipment
5. Provide a sufficiently low-resistance path to ground to minimize rise in ground potential with respect to remote ground

Substation safety requirements call for the grounding of all exposed metal parts of switches, structures, transformer tanks, metal walkways, fences, steelwork of buildings, switchboards, instrument-transformer secondaries, etc. so that a person touching or near any of this equipment cannot receive a dangerous shock if a high-tension conductor flashes to or comes in contact with any of the equipment listed. This function in general is satisfied if all metalwork between which a person can complete contact or which a person can touch when standing on the ground is so bonded and grounded that dangerous potentials cannot exist. This means that each individual piece of equipment, each structural column, etc., must have its own connection to the station grounding mat.

A most useful source of information concerning substation grounding is contained in the comprehensive guide IEEE Standard 80-1986, IEEE Guide for Safety in AC Substation Grounding Period. Much of the following information is based on recommendations stated in the IEEE Standard 80.

The basic substation ground system used by most utilities takes the form of a grid of horizontally buried conductors. The reason that the grid or mat is so effective is attributed to the following:

1. In systems where the maximum ground current may be very high, it is seldom possible to obtain a ground resistance so low as to ensure that the total rise of the grounding system potential will not reach values unsafe for human contact. This being the case, the hazard can be corrected only by control of local potentials. A grid is usually the most practical way to do this.

2. In HV and EHV substations, no ordinary single electrode is adequate to provide needed conductivity and current-carrying capacity. However, when several are connected to each other and to structures, equipment frames, and circuit neutrals which are to be grounded, the result is necessarily a grid, regardless of original objectives. If this grounding network is buried in soil of reasonably good conductivity, this network provides an excellent grounding system.

The first step in the practical design of a grid or mat consists of inspecting the layout plan of equipment and structures. A continuous cable should surround the grid perimeter to enclose as much ground as practical and to avoid current concentration and hence high gradients at projecting ground cable ends. Within the grid, cables should be laid in parallel lines and at reasonably uniform spacing. They should be located, where practical, along rows of structures or equipment to facilitate the making of ground connections. The preliminary design should be adjusted so that the total length of buried conductor, including cross connections and rods, is at least equal to that required to keep local potential differences within acceptable limits.

A typical grid system for a substation might comprise 4/0 bare stranded copper cable buried 12 to 18 in below grade and spaced in a grid pattern of about 10 by 20 ft. (Other conductor sizes, burial depths, and grid conductor spacings, however, are frequently used.) At each junction of 4/0 cable, the cables would be securely bonded together, and there might also be connected a driven copper-covered steel rod approximately $^5/_8$ in. in diameter and approximately 8 ft long. In very high-resistance soils it might be desirable to drive the rods deeper. (Lengths approaching 100 ft are recorded.) A typical grid system usually extends over the entire substation yard and sometimes a few feet beyond the fence, which surrounds the building and equipment. Figure 17-19 shows a grounding plan for a typical EHV substation operating at 345 kV.

In order to ensure that all ground potentials around the station are equalized, the various ground cables or buses in the yard and in the substation building should be bonded together by heavy multiple connections and tied into the main station ground. This is necessary in order that appreciable voltage differences to ground may not exist between the ends of cables which may run from the switchyard to the substation building.

Heavy ground currents, such as those that may flow in a transformer neutral during ground faults, should not be localized in ground connections (mats or groups of rods) of small area in order to minimize potential gradients in the area around the ground connections. Such areas should have reinforced wire sizes where necessary to handle adequately the most severe condition of fault-current magnitude and duration.

Copper cables or straps are usually employed for equipment-frame ground connections. However, transformer tanks are sometimes used as part of the ground path for lightning arresters mounted thereon. Similarly, steel structures may be used as part of the path to ground if it can be established that the conductivity, including that of any joints, is and can be maintained as equivalent

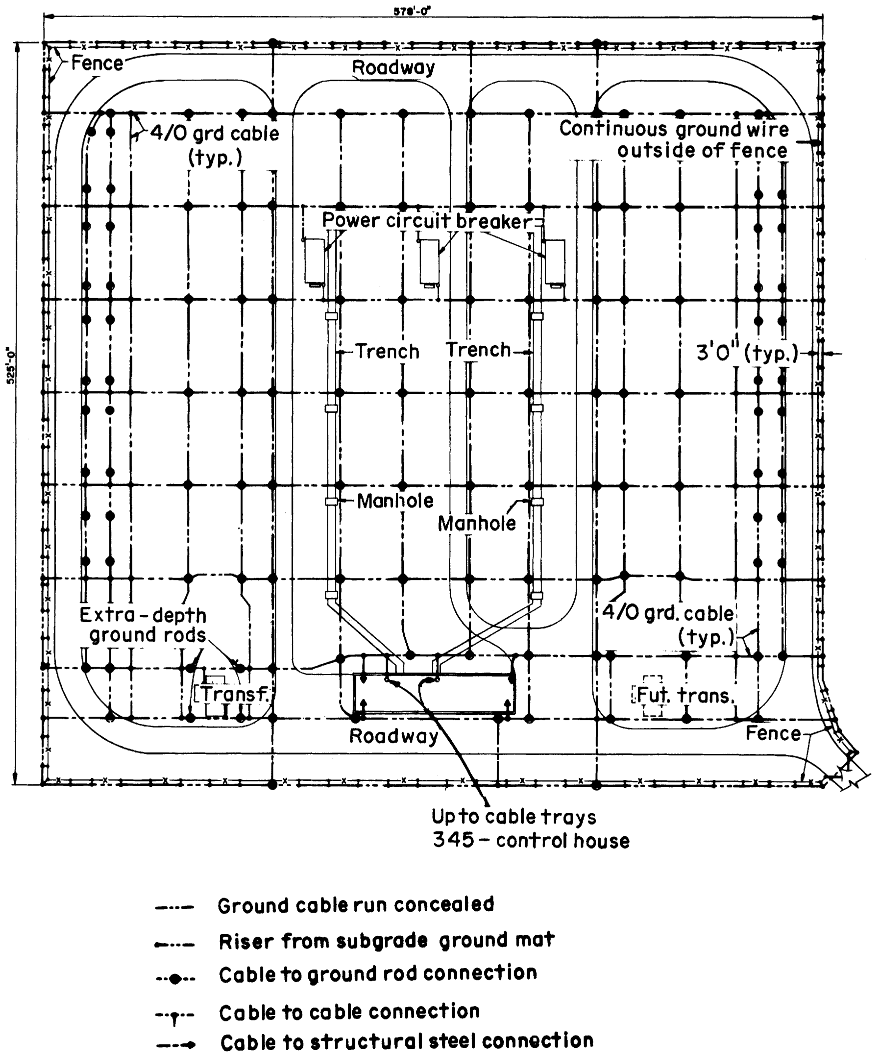


FIGURE 17-19 Grounding plan for a 345-kV substation.

to the copper conductor that would otherwise be required. Studies by some utilities have led to their successful use of steel structures as part of the path to the ground mat from overhead ground wires, lightning arresters, etc. Where this practice is followed, any paint films, which might otherwise introduce a high-resistance joint should be removed and a suitable joint compound applied or other effective means taken to prevent subsequent deterioration of the joint from oxidation.

Connections between the various ground leads and the cable grid and connections within the cable grid are usually clamped, welded, or brazed. Ordinary soldered connections are to be avoided because of possible failure under high fault currents or because of galvanic corrosion.

Each element of the ground system (including grid proper, connecting ground leads, and electrodes) should be so designed that it will

1. Resist fusing and deterioration of electric joints under the most adverse combination of fault-current magnitude and fault duration to which it might be subjected.
2. Be mechanically rugged to a high degree, especially in locations exposed to physical damage.
3. Have sufficient conductivity so that it will not contribute substantially to dangerous local potential differences.

Adequacy of a copper conductor and its joints against fusing can be determined from Table 17-10 and by referring to Fig. 17-20.

If the switchyard is on soil of high resistivity so that it is impossible to obtain suitably low resistance from rods driven within the station, it is possible to reduce the resistance by extending the main ground grid outside the enclosed substation area to a secondary ground mat located adjacent to the substation. The effective resistance of the complete grounding system can be lowered appreciably by the use of a more extensive grid area and of additional grid conductor length. An important reason for trying to lower grid resistance is to minimize ground-potential rise with respect to remote ground during ground faults.

TABLE 17-10 Minimum Copper Conductor Sizes to Avoid Fusing

	Circular mils per ampere		
Time duration of fault, s	Cable only	With brazed joints	With bolted joints
30	40	50	65
4	14	20	24
1	7	10	12
0.5	5	6.5	8.5

Ground-potential rise depends on fault-current magnitude, system voltage, and ground-system resistance. The current through the ground system multiplied by its resistance measured from a point remote from the substation determines the ground-potential rise with respect to remote ground. The current through the grid is usually considered to be the maximum available line-to-ground fault current. For example, a ground fault of 15,000 A flowing into a ground grid with a value of 0.5 Ω resistance to absolute earth would cause an *IR* drop of 7500 V. The 7500-V *IR* drop due to the fault current could cause serious trouble to communications lines entering the station if the communications facilities are not properly insulated or neutralized.

Low-resistance station grounds are frequently difficult to obtain. In such cases, the use of driven grounds will provide the most convenient means of obtaining a suitable ground connection. The arrangement and number of driven grounds will depend on the station size and the nature of the soil. The ground mat of Fig. 17-19 has a value measured to be on the order of 0.5 Ω. The best soils for ground mats are wet and marshy, with clay or clay loam as the next best. Sand and sandy soils are of higher resistance, making it difficult to obtain low-resistance ground connections.

The size of the rods used is determined mainly by the depth to which they must be driven, although small rods can be driven to considerable depths by the use of driving collars. Figure 17-21 shows the relationship between rod size and resistance obtained. Driving more rods in a given space will help reduce resistance, but the reduced resistance

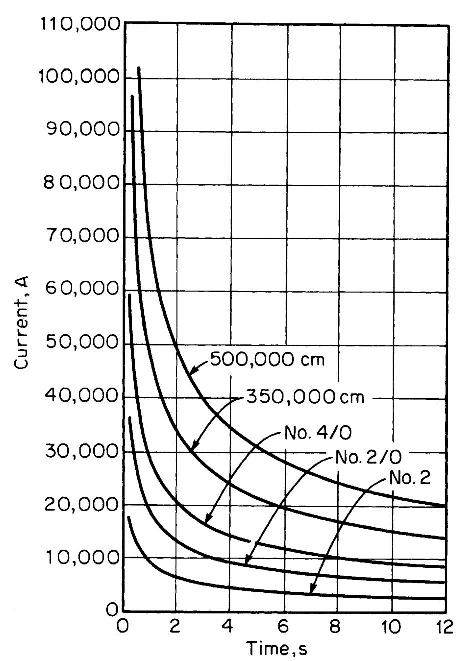


FIGURE 17-20 Short-time fusing curves for copper cable.

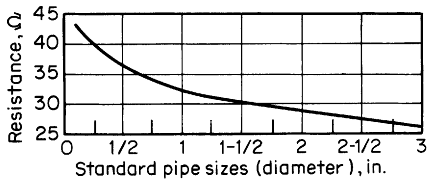


FIGURE 17-21 Relation between pipe diameter and ground resistance. *(NBS Technologic Paper No. 108, June 1918.)*

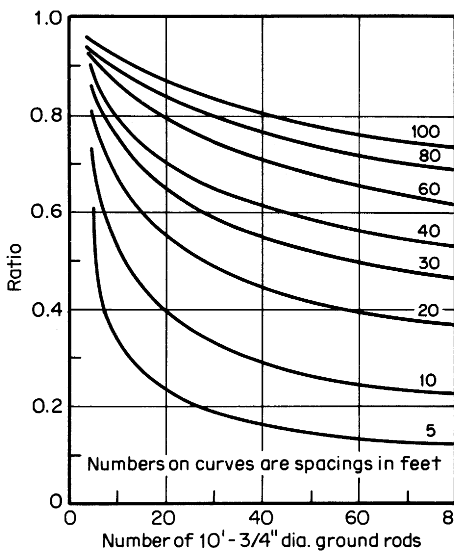


FIGURE 17-22 Ratio of conductivity of ground rods in parallel on an area to that of isolated rods. *(H. B. Dwight, Trans. AIEE, vol. 55, p. 1936.)*

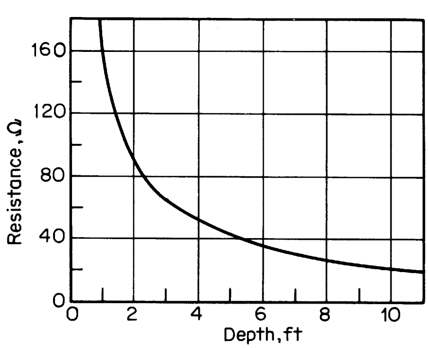


FIGURE 17-23 Variation of resistance of driven pipes with depth. Soil fairly wet. External diameter of pipe is 24.9 mm (1.02 in). *(NBS Technologic Paper No. 108, June 1918.)*

is not a function of the number of rods. Figure 17-22 shows the effect on resistance of spacing and number of rods in square areas. These curves apply to $^3/_4$-in by 10-ft rods. The rods or pipes for permanent stations should be of noncorroding materials. Figure 17-23 shows the effect of increased length of rods in uniform soil. Usually the improvement is much greater than indicated because the rods penetrate into better-conducting earth as they are driven deeper. In addition, where the ground can become frozen, rods must be driven below the frost line to obtain low resistance.

In general, it is advisable to obtain reduced ground resistance by the use of a more extensive mat and more ground rods rather than by treating the earth around the rods with salt because of the impermanence of the treatment. However, treatment of the soil is sometimes the only means whereby suitable resistance can be obtained.

It is not possible to describe all methods of obtaining ground connections of suitably low resistance. The problem sometimes presents great difficulties and calls for considerable extra expense. Substations should not be located on solid rock with little or no topsoil, since the cost of obtaining a low-resistance ground would be excessive. Such a ground would require the use of an extensive counterpoise system with many drilled "wells," in which electrodes would be inserted in treated filling, with provision made for renewing the treatment.

Measuring Ground Resistance. The measurement of ground resistance is necessary both at the time of initial energization of a substation and at periodic intervals thereafter to ensure that the value of ground resistance does not increase appreciably. The measurement of the resistance of a ground connection with respect to absolute earth is somewhat difficult. All results are approximations and require careful application of the test equipment and selection of reference ground points.

There are several methods of testing ground resistance, but all of them are similar in that two reference ground connections are used and a suitable source of current is required for the test. Some form of alternating current is circulated through the ground under test in amounts from a few milliamperes, as in bridge methods and with some of the patented ground testers, up to 100 A or more. The amount of current used depends on the method, and methods using very small currents will give results as accurate as methods using heavy currents if the ground under test is one for which the test method is suitable.

Methods of testing ground resistance fall into three general groups:

1. *Triangulation or three-point methods*, in which two auxiliary test grounds and the point to be measured are arranged in a triangular configuration. The series resistance of each pair of ground points in the triangle is determined by measuring the voltage across and the current through the ground resistance being measured. Resistance measurements are made by the voltmeter-ammeter method or by means of a suitable bridge. For accurate results, the resistance of the auxiliary grounds and the ground under test should be of the same order of magnitude, and results may be meaningless if the test grounds have more than 10 times the resistance of the ground under test. This method is suitable for measuring the resistance of tower footings, isolated ground rods, or small grounding installations. It is not suitable for measurement of low-resistance grounds such as the ground grid at large substations.
2. *Ratio methods,* in which the series resistance R of the ground under test and a test probe is measured by means of a bridge which operates on the null-balance principle. A calibrated slide-wire potentiometer is connected to the two ground connections, with the slider of the potentiometer connected to a second test probe. The potential of the slider to ground is adjusted to zero or null. If D is the total slide-wire resistance and d_1 is the resistance from the slider to the ground under test, the resistance R of the ground under test is $(d_1/D) \times R$. The vibrometer and the groundometer, self-contained test instruments, make use of this principle. This method is much more satisfactory than triangulation methods, since ratios of test-probe resistance to the resistance of the ground under test run as high as 300:1 with test instruments such as the groundometer. Although this method has its limitations in testing low-resistance grounds of large areas, suitable readings can be obtained by locating the test probes in a straight line, in a direction 90° from the substation fence, and with the distance of the farthest probe twice the width of the substation. Best accuracy can be attained by taking measurements at the greatest possible distance from the ground grid being measured.
3. *Fall-of-potential methods,* which include methods using close-in reference grounds, usually less than 1000 ft from the ground under test. The principle of the fall-of-potential method using close-in reference grounds is illustrated by Fig. 17-24. A fixed probe is driven in the ground at point C_2 with a movable probe P_2 set at various points in a straight line between C_2 and the ground mat G

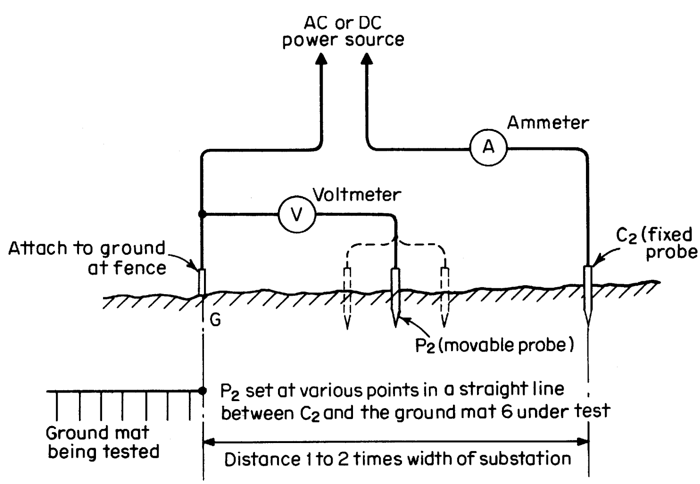


FIGURE 17-24 Field setup for making ground-resistance tests by means of the fall-of-potential method.

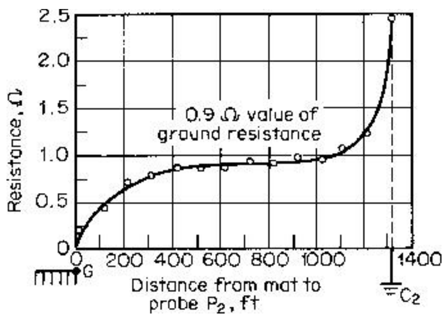


FIGURE 17-25 Ground-resistance curve for a substation ground mat.

under test. Either alternating current or direct current is circulated through ground *G* and fixed test probe C_2. A voltmeter is connected between point *G* and probe P_2, and an ammeter is connected to observe current flow through probe C_2. Voltmeter readings *E* are taken simultaneously with ammeter readings *I*. The reading *E*/*I*, which equals the resistance in ohms, is plotted in Fig. 17-25. The resistance shown on the flat part of the curve or at the point of inflection is taken as the resistance of the ground. This method may be subject to considerable error if stray currents are present. It is normally applied by using several test-probe readings at 10% intervals of the distance from *G* to C_2, with the test probe located midway between *G* and C_2. Self-contained test instruments which make use of this method are available; among them are the ground ohmer and Megger ground tester. These instruments give better results than the voltmeter-ammeter method, since they are designed to eliminate the effects of stray currents.

In recent years, considerable emphasis has been placed on the use of computer programs to calculate the design parameters of substation ground systems. These programs normally employ methods detailed in IEEE Guide 80. Normal input data required to run a typical program consist of the following:

1. System voltage, symmetric rms single-phase-to-ground fault currents, and the clearing time of faults
2. Length and width of substation area
3. Estimated value of ground resistivity in ohm-meters
4. Assumed value of ground conductor length
5. Cross section of conductors available

The following typical information is derived from the program:

1. Size, total length, and number of strands of copper ground conductor
2. Spacing of main grid configuration along width and length
3. Expected ground-mat resistance
4. Depth of grid below ground level
5. Tolerable limits and maximum values of step and touch potentials

It should be noted that the step and touch potentials are defined as follows:

E_{step} is the tolerable potential difference between any two points on the ground surface, which can be touched simultaneously by the feet.

E_{touch} is the tolerable potential difference between any point on the ground where a person may stand and any point which can be touched simultaneously with either hand.

17.1.11 Transformers

Transformers Connections. *Delta-delta–connected* transformers are used mainly on the lower transmission voltages. This is due to the fact that the complete winding must be insulated for full line-to-line voltage; for voltages above 73 kV, the cost increase is appreciable over Y-connected transformers with graded insulation. The delta-delta–connected transformers have one advantage in that the bank can be operated open delta at 86.6% of the capacity of the two remaining transformers.

The delta-star connection is in common use for both step-up and step-down purposes. When used as a step-up transformer, the high-tension winding is Y-connected; when used for step-down purposes, the low-tension winding is usually Y-connected in order to provide a grounded neutral for secondary transmission or for primary distribution.

Delta-connected high-tension windings, however, are seldom used for transmission voltages of 138 kV and above. The delta-Y connection almost completely suppresses the triple harmonics with the neutral solidly grounded. Triple harmonics, which can appear on power systems are the third and its odd multiples. Y-connected windings on the higher voltages are usually provided with graded insulation, the neutral-end turns of which may have very little insulation if the neutral is solidly grounded. If neutral impedance (reactor or resistor) is used, neutral insulation must be equal to or greater than the maximum *IZ* drop of the neutral impedance. If the neutral is to be left ungrounded on either grounded neutral systems or ungrounded neutral systems, the neutral insulation should be the same as it is on the line side to avoid traveling-wave troubles.

Star-star–connected (Y-Y) transformers are used infrequently on high-voltage transmission systems. When used with both neutrals grounded, if single-phase or three-phase shell type, they must be used with Y-connected generators, and a solid neutral connection must be provided between the generator, or generators, and the low-tension transformer neutral in order to minimize triple-harmonic troubles. The various types of Y-Y–connected transformers can be used with both neutrals ungrounded with satisfactory results or with neutrals grounded if of the three-phase core type. The triple harmonics are nearly suppressed in three-phase core-type transformers.

Star-star–connected transformers with a delta-connected third winding (tertiary) overcome the difficulties of the simple Y-Y connection. The tertiary winding may be for the suppression of harmonics only, in which case no connections are brought out with three-phase transformers. Y-delta-Y transformers are frequently used to supply two distribution voltages or a distribution voltage and a secondary-transmission voltage. If the service supplied from the delta-connected winding is four-wire three-phase, the neutral must be obtained from a separate grounding transformer. A common use for the tertiary winding is to provide substation station-service power to operate station auxiliary equipment. Three-winding transformers, all windings of which are used, are frequently rated with two outputs: (1) the individual output of each secondary winding alone with the other secondary winding carrying no load and (2) a simultaneous loading rating in which each secondary winding is given a rated loading with the primary-winding loading the resultant of the two secondary loadings.

Autotransformers are generally used for transforming from one transmission voltage to another when the ratio is 3:1 or less. Such transformers are normally connected in Y with the neutral solidly grounded and when so connected should be provided with a closed delta tertiary winding of adequate capacity for the suppression of harmonics, for ground-fault duty, and to provide station-service power. The tertiary is frequently used to provide a supply of distribution voltage. Autotransformers are superior to separate-winding transformers owing to their lower cost, greater efficiency, smaller size and weight, and better regulation. Autotransformers also may be obtained with zigzag-connected windings or with delta-connected windings. Both these types are free from triple-harmonic troubles but in general are more expensive.

Delta-connected autotransformers have a possible disadvantage in that they insert a phase shift into the transformation, which means that the system being served must be radial or else it must be served by similar transformations at other points.

Transformer Loading Practice. Because of the varying load cycle of most transformers, it is customary to permit loading considerably in excess of the transformer nameplate rating. There may be limitations on the transformer imposed by bushings, leads, tap changers, cables, disconnecting switches, circuit breakers, etc. Good engineering design, however, will permit operation without these limitations.

The increase in transformer loading is limited by the effect of temperature on insulation life. High temperature decreases the mechanical strength and increases the brittleness of fibrous insulation and makes transformer failure increasingly likely even though the dielectric strength of the insulation may not be seriously decreased. Overloading should be limited then by giving consideration to the

TABLE 17-11 Percent Daily Peak Load for Normal Life Expectancy with 30°C Cooling Air

Duration of peak load, h	Self-cooled with % load before peak of			Forced-air-cooled up to 133% of self-cooled rating, with % load before peak of			Forced-air-cooled over 133% of self-cooled rating, or forced-oil-cooled, with % load before peak of		
	50%	70%	90%	50%	70%	90%	50%	70%	90%
0.5	189	178	164	182	174	161	165	158	150
1	158	149	139	150	143	135	138	133	128
2	137	132	124	129	126	121	122	119	117
4	119	117	113	115	113	111	111	110	109
8	108	107	106	107	107	106	106	106	105

effect on insulation life and transformer life. For recurring loads, such as the daily load cycles, the transformer would be operated for normal life expectancy. For emergencies, either planned or accidental, loading would be based on some percentage loss of life.

In a typical case for a failure of part of the electrical system, a 2.5% loss of life per day for a transformer may be acceptable. Loading recommendations based on the evaluation of the loss of insulation life as affected by temperature are contained in ANSI Standard C57.91-1995, Institute of Electrical and Electronics Engineers Guide for Loading Mineral-Oil-Immersed Transformers. NEMA Publ. TR98-1964 contains corresponding recommendations for loading power transformers with 65°C average winding rise insulation systems. ANSI Standard C57.91-1995 states that an average loss of life of 1% per year or 5% in any one emergency operation is considered reasonable.

Daily overload cycles consistent with normal life expectancy for air-cooled power transformers at 30°C ambient temperature are given in Table 17-11, which is a condensation of data taken from ANSI Standard C57.91-1995. For a listing of transformer loading above normal with some sacrifice of life expectancy, data given in NEMA Publ. TR98-1964, Part 3, are condensed in Table 17-12.

Ambient temperature affects load capacity by an amount depending on the type of cooling as shown in Tables 17-11 and 17-12. For changes from this average ambient temperature, transformer ratings may be adjusted as shown in Table 17-13. The table applies to both the 55°C and the 65°C average winding-temperature-rise transformers. For the ambient temperature of air-cooled transformers, use the average value over a 24-h period or 10°C under the maximum during the 24-h period, whichever is higher.

The following temperatures and load limitations are generally applied to transformers. The temperature of the top oil should never exceed 100°C. The maximum hot-spot winding temperature should not exceed 150°C for 55°C rise transformers or 180°C for 65°C rise transformers. Short-time peak loading for $^1/_2$ h or more should not exceed 200% rating. At abnormally high temperatures it may be necessary to remove some oil in order to avoid overflow or excessive pressure.

17.1.12 Surge Protection

A substation should be designed to include safeguards against the hazards of abnormally high voltage surges that can appear across the insulation of electrical equipment in the station. The most severe overvoltages are caused by lightning strokes and by switching surges. The main methods to prevent these overvoltages from causing insulation failures include:

1. Use of surge arresters
2. Equipment neutral grounding
3. Proper selection of equipment impulse insulation level

TABLE 17-12 Allowable Peak Loads (in Multiples of Maximum Nameplate Rating) for Moderate Sacrifice of Life Expectancy with 30°C Cooling Air

Duration of peak load, h	Hottest-spot temperature reached, °C	Life loss in percent not more than	Self-cooled (OA) with % load before peak of				Forced-air-cooled (OA/FA) up to 133% of self-cooled rating with % load before peak of				Forced-air-cooled (OA/FA/FA) over 133% of self-cooled rating or forced-oil-cooled (FOA or OA/FOA/FOA) with % load before peak of			
			50%	70%	90%	100%	50%	70%	90%	100%	50%	70%	90%	100%
½	171	0.25	2.00	2.00	2.00	1.96	2.00	1.95	1.85	1.80	1.64	1.60	1.54	1.51
	180	0.50	2.00	2.00	2.00	2.00	2.00	2.00	1.95	1.90	1.69	1.66	1.60	1.57
1	163	0.25	1.96	1.89	1.80	1.74	1.77	1.72	1.65	1.61	1.47	1.45	1.49	1.39
	180	1.00	2.00	2.00	1.99	1.94	1.93	1.88	1.81	1.78	1.57	1.55	1.52	1.50
2	155	0.25	1.68	1.63	1.57	1.53	1.53	1.50	1.47	1.44	1.33	1.32	1.31	1.30
	171	1.00	1.83	1.79	1.71	1.64	1.66	1.64	1.60	1.58	1.42	1.41	1.39	1.39
	180	2.00	1.91	1.83	1.71	1.64	1.74	1.71	1.65	1.61	1.47	1.46	1.44	1.43
4	147	0.25	1.44	1.41	1.39	1.37	1.35	1.34	1.33	1.32	1.24	1.23	1.23	1.23
	163	1.00	1.55	1.52	1.47	1.44	1.47	1.46	1.45	1.45	1.32	1.32	1.32	1.32
	180	4.00	1.55	1.52	1.47	1.44	1.51	1.50	1.47	1.46	1.40	1.40	1.39	1.39
8	139	0.25	1.28	1.27	1.27	1.26	1.24	1.24	1.24	1.24	1.18	1.18	1.18	1.18
	155	1.00	1.38	1.37	1.36	1.36	1.36	1.36	1.36	1.36	1.27	1.27	1.27	1.27
	171	4.00	1.38	1.37	1.36	1.36	1.42	1.42	1.41	1.41	1.35	1.35	1.35	1.35

Note: For forced-air-cooled transformers, the peak loads are calculated on the basis of all cooling being in use during the period preceding the peak load. When operating without fans, use the tables for OA transformers. Differences in cooling methods used with forced-oil-cooled transformers result in differences in peak-load-carrying ability. Consult the manufacturer before applying loads above the values given in the table.

Source: Based on capability tables in *NEMA Publ.* TR98, Part 3.

TABLE 17-13 Effect of Ambient Temperature on kVA Capacity

Type of cooling	% of rated kVA decrease in capacity for each °C increase over 30°C air	% of rated kVA increase in capacity for each °C decrease under 30°C
Self-cooled—OA	1.5	1.0
Forced-air-cooled—OA/FA, OA/FA/FA	1.0	0.75
Forced-air-cooled—FOA, OA/FOA/FOA	1.0	0.75

4. Proper selection and coordination of equipment basic insulation levels

5. Careful study of switching-surge levels that can appear in the substation

The main device used to prevent dangerous overvoltages, flashovers, and serious damage to equipment is the surge arrester. The surge arrester conducts high surge currents, such as can be caused by a lightning stroke, harmlessly to ground and thus prevents excessive overvoltages from appearing across equipment insulation. For a detailed description of the characteristics and application of arresters, refer to Sec. 27.

The important consideration in applying surge arresters and in selecting equipment insulation levels depends greatly on the method of grounding used. Systems are considered to be effectively grounded when the coefficient of grounding does not exceed 80%. Similarly, systems are noneffectively grounded or ungrounded when the coefficient of grounding exceeds 80%.

A value not exceeding 80% is obtained approximately when, for all system conditions, the ratio of zero sequence reactance to positive sequence reactance (X_0/X_1) is positive and less than 3 and the ratio of zero sequence resistance to positive sequence reactance (R_0/X_1) is positive and less than 1. What this says in effect is that if neutrals are grounded solidly everywhere and if a ground occurs on one of the conductors, then the voltage that can appear on the healthy phases cannot exceed 80% of normal phase-to-phase voltage.

Thus, the *coefficient of grounding* is defined as the ratio of maximum sustained line-to-ground voltage during faults to the maximum operating line-to-line voltage. On many HV and EHV systems, the coefficient of grounding may be as low as 70%.

Surge-arrester ratings are normally selected on the basis of the coefficient of grounding; thus, for effectively grounded systems, the 80% arrester is selected when using the conventional gap-type arrester. When using the gapless metal oxide arrester, a lower-value arrester may be selected based on the maximum continuous operating voltage (MCOV) equal to the maximum normal line-to-neutral voltage. For example, a 115-kV system (maximum operating voltage equals 121 kV) can use a 97-kV conventional arrester, that is, 80% of 121 kV, when operating on a solidly grounded system, and can use a gapless-type metal oxide arrester rated 70 kV. It should be noted that other factors, such as resonant conditions and system switching, could increase the value of the coefficient of grounding and thus should be studied in each individual system.

The impulse insulation level of a piece of equipment is a measure of its ability to withstand impulse voltage. It is the crest value, in kilovolts, of the wave of impulse voltage that the equipment must withstand. However, at EHV, the switching-surge insulation level may be lower than the corresponding impulse level, and thus the switching-surge level becomes the dominant factor in establishing insulation levels.

Basically, the coordination of insulation in a substation means the use of no higher-rated arrester than required to withstand the 60-Hz voltage and the choice of equipment insulation levels that can be protected by the arrester. Careful study of switching-surge levels that can occur at the substation as determined, for example, by transient network analyzer studies also can be used to determine and coordinate proper impulse insulation and switching-surge strength required in a substation electrical equipment.

REFERENCES ON AIR-INSULATED SUBSTATIONS

Books

Alcoa Aluminum Bus Conductor Handbook. Pittsburgh, Pa., Aluminum Company of America, 1957.

Applied Protective Relaying. Newark, N.J., Westinghouse Electric Corporation.

Blume, L. F., Boyajian, A., Camilli, G., Lennox, T. S., Minneci, S., and Montsinger, V. M.: *Transformer Engineering.* New York, John Wiley & Sons, 1951.

Electrical Transmission and Distribution Reference Book. East Pittsburgh, Pa., Westinghouse Electric Corporation, 1950.

Mason, C. R.: *The Art and Science of Protective Relaying.* New York, John Wiley & Sons, 1967.

Van C. Warrington, A. R.: *Protective Relays, Their Theory and Practice,* vol. 1. London, Chapman & Hall, 1971.

Van C. Warrington, A. R.: *Protective Relays, Their Theory and Practice,* vol. 2. London, Chapman & Hall, 1974.

Station Design

AIEE Committee Report: A Guide for Minimum Electrical Clearances for Standard Basic Insulation Levels, *Trans. AIEE, Power Appar. Syst.,* June 1954, vol. 73, pp. 636–641.

AIEE Committee Report: Basic Structural Design for Transmission Substations Including Light Metals, *Electr. Eng.,* April 1952, vol. 71, pp. 344–350.

Colombo, A., Sartorio, G., and Taschini, A.: Phase to Phase Air Clearances in EHV Substations as Required by Switching Surges, *CIGRE Paper* 33-11, 1972.

Committee Report: Design Standardization Methods and Techniques for Substation Facilities (Bibliogr.), *Trans. AIEE, Power Appar. Syst.,* October 1964, vol. 83, pp. 1029–1034.

Committee Report: 500 kV AC Substation Design Criteria, Summary of Industry Practices, *IEEE Trans., Power Appar. Syst.,* 1969, vol. 88, pp. 854–861.

Committee Report: Minimum Line-to-Ground Electrical Clearances for EHV Substations Based on Switching Surge Requirement, *IEEE Trans., Power Appar. Syst.,* 1972, vol. 91, pp. 1924–1930.

Committee Report: 700/765 kV AC Substation Design Criteria, A Summary of Industry Practices, *IEEE Trans., Power Appar. Syst.,* 1970, vol. 89, pp. 1521–1524.

Dolan, P. R., and Peat, A. J.: Design of the First 500 kV Substations on the Southern California Edison Company System, *IEEE Trans., Power Appar. Syst.,* 1967, vol. 86, pp. 531–539.

Hertig, G. E.: High- and Extra-High-Voltage Substation Design and Economic Comparisons, *Trans. AIEE,* 1963, vol. 81, pp. 832–840.

IEEE Standard C2-1997, National Electrical Safety Code.

Paris, L., and Taschini, A.: Phase-to-Ground and Phase-to-Phase Air Clearances in Substations, *CIGRE, Electra,* 1973, no. 29. (Recommended by CIGRE S.C. 23 and CIGRE S.C. 33.)

Scherer, H. N.: 765 kV Station Design, *IEEE Trans. Power Appar. Syst.,* 1969, vol. 88, pp. 1372–1376.

Bus Construction

Attri, N. S., and Edgar, J. N.: Response of Bus Bars on Elastic Supports Subjected to a Suddenly Applied Force, *IEEE Trans., Power Appar. Syst.,* 1967, vol. 86, pp. 636–650.

Committee Report: Use of Aluminum for Substation Busses, *IEEE Trans., Power Appar. Syst.,* 1963, vol. 82, pp. 72–102.

Dwight, H. B.: Skin Effect and Proximity Effect in Tubular Conductors, *Trans. AIEE,* February 1922, pp. 189–198.

Fischer, E. G.: Seismic Design of Bus Runs and Supports, *IEEE Trans., Power Appar. Syst*., 1973, vol. 92, pp. 1493–1500.

Foti, A.: Design and Application of EHV Disconnecting Switches, *Trans. AIEE, Power Appar. Syst.,* October 1965, vol. 84, pp. 868–876.

Higgins, T. J.: Formulas for Calculating Short Circuit Forces between Conductors of Structural Shape, *Trans. AIEE,* October 1943, vol. 62, pp. 659–663.

Higgins, T. J.: Formulas for Calculating Short Circuit Stresses for Bus Supports for Rectangular Tubular Conductors, *Trans. AIEE,* August 1942, vol. 61, pp. 578–580.

IEEE Standard C37.30-1992, Requirements for High Voltage Air Switches.

IEEE Standard C37.32-1992 (American National Standard for Switchgear, High-Voltage Air Switches, Bus Supports, and Switch Accessories), Schedules of Preferred Ratings, Manufacturing Specifications, and Application Guide.

Milton, R. M., and Chambers, F.: Behavior of High-Voltage Busses and Insulators during Short Circuits, *Trans. AIEE,* August 1955, vol. 74, pp. 742–749.

NEMA Standard SG6-1995, Power Switching Equipment.

Rayleigh, J. W. S.: *Aeolian Tones.* New York, Cambridge University Press, 1920.

Schurig, O. R., and Sayre, M. F.: Mechanical Stresses in Bus Bar Supports During Short Circuits, *J. AIEE,* April 1925, vol. 44, pp. 365–372.

Taylor, D. W., and Stuehler, C. M.: Short Circuit Tests on 138 kV Busses, *Trans. AIEE,* August 1956, vol. 75, pp. 739–747.

Temple, G., and Brickley, W. G.: *Rayleigh's Principle.* New York, Oxford University Press, 1933.

Wagner, C. F.: Current Distribution in Multi Conductor Single Phase Buses, *Electr. World,* March 18, 1922, vol. 79, no. 11.

System Protection

Blackburn, J. L.: Future Automatic Switching of EHV Transmission Lines—Development and Application of Solid-State Relays, *Proc. Am. Power Conf.,* 1965, vol. 27, pp. 998–1008.

Boyaris, E., and Guyot, W. S.: Experience with Fault Pressure Relaying and Combustible Gas Detection in Power Transformers, *Proc. Am. Power Conf.,* 1971, vol. 33, pp. 1116–1126.

Chadwick, J. W., and Goff, L. E.: Development of a Static Single-pole Relaying Scheme for the TVA 500-kV System, *Proc. Am. Power Conf.,* 1971, vol. 33, pp. 1127–1133.

Committee Report: Relaying the Keystone 500 kV System, *IEEE Trans.,* June 1968, vol. 87, no. 5, no. 6, pp. 1434–1439.

Elmore, W. A.: Some Guidelines for Selecting a Solid-State Transmission Line Relaying System, *Westinghouse Eng.,* March 1972, vol. 32, no. 2, pp. 50–59.

Emanuel, A. E., and Vora, J. P.: Sensor Coil for Internal Fault Protection of Shunt Reactors, *IEEE Trans.,* November-December 1974, vol. 93, no. 6, pp. 1917–1926.

Forford, T., and Linders, J. R.: A Half Cycle Bus Differential Relay and Its Applications, *IEEE Trans.,* July-August 1974, vol. 93, no. 4, pp. 1110–1120.

Horowitz, S. H., and Seeley, H. T.: Relaying the AEP 765 kV System, *IEEE Trans.,* September 1969, vol. PAS-88, no. 9, pp. 1382–1389.

IEEE Committee Report: Bibliography of Relay Literature 1995.

IEEE Committee Report: Ground Relaying Practices and Problems: A Power System Relaying Committee Survey, *IEEE Trans., Power Appar. Syst.,* May 1966, vol. PAS-85, no. 5, pp. 524–532.

IEEE Standard C37.90-1-1989, Guide for Surge Withstand Capability (SWC) Tests.

IEEE Standard C37.90-1989, Relays and Relay Systems Associated with Electric Power Apparatus.

IEEE Standard C37.91-1985, Guide for Protective Relay Applications to Power Transformers.

Korponay, N., and Ungrad, H.: The Requirements Made of Current Transformers by High-Speed Protective Relays, *Brown Boveri Rev.,* June 1968, vol. 55, no. 6, pp. 289–297.

Narayan, V.: Distance Protection of H. V. and E. H. V. Transmission Lines, *Brown Boveri Rev.,* July 1971, vol. 58, no. 7, pp. 276–286.

Rockefeller, G. D.: What Are the Prospects for Substation-Computer Relaying? *Westinghouse Eng.,* September 1972, vol. 32, no. 5, pp. 152–156.

Schumm, G. P.: The Philosophy of Protective Relaying in the United States and Europe, *Proc. Am. Power Conf.,* 1971, vol. 33, pp. 1105–1115.

Sutton, H. J.: The Application of Relaying on an EHV System, *IEEE Trans.,* April 1967, vol. 86, no. 4, pp. 408–415.

Sykes, J. A., and Morrison, I. F.: A Proposed Method of Harmonic Restraint Differential Protection of Transformers by Digital Computer, *IEEE Trans.,* May-June 1972, vol. 91, no. 3, pp. 1266–1272.

Ungrad, H.: Back-up Protection, *Brown Boveri Rev.,* June 1968, vol. 55, no. 6, pp. 297–305.

Ungrad, H.: Distance Relays with Signal Transmission for Main and Backup Protection, *Brown Boveri Rev.,* July 1971, vol. 58, no. 7, pp. 293–304.

Vanderleck, J. M.: Measurement of Composite Error of Relay-type Current Transformers, *Ontario Hydro Res. Q.,* 1967, vol. 19, no. 3, pp. 15–18.

Shielding of Control Cables

Dietrick, R. E., Ramberg, H. C., and Barber, J. C.: BPA Experience with EMI Measurements and Shielding in EHV Substations, *Proc. Am. Power Conf.,* 1970, vol. 32, pp. 1054–1061.

Kotheimer, W. C.: Control Circuit Transients, Pt. 1, *Power Eng.,* January 1969, vol. 73, no. 1, pp. 42–46.

Kotheimer, W. C.: Control Circuit Transients, Pt. 2, *Power Eng.,* February 1969, vol. 73, no. 2, pp. 54–56.

Sutton, H. J.: Transients Induced in Control Cables Located in EHV Substation, *IEEE Trans.,* July-August 1970, vol. 89, no. 6, p. 1069.

Grounding

AIEE Committee Report: Application Guide on Methods of Substation Grounding, *Trans. AIEE, Power Appar. Syst.*, April 1954, vol. 73, pp. 271–275.

Bellasi, P. L.: Impulse and 60-Cycle Characteristics of Driven Grounds, *Trans. AIEE,* March 1941, vol. 60, pp. 123–128.

Committee Report: Principles and Practices in Grounding, *Edison Electr. Inst. Ser. Rep.* D9, October 1936.

Eaton, J. R.: Grounding Electric Circuits Effectively, I, II, III, *Gen. Electr. Rev.,* June, July, and August 1941.

EPRI Final Report EL-2682: Analysis Techniques for Power Substation Grounding Systems, vol. 1, Design Methodology and Tests.

IEEE Standard 80-1986, IEEE Guide for Safety in AC Substation Grounding.

IEEE Standard 81-1983, Recommended Guide for Measuring Ground Resistance and Potential Gradients in the Earth.

IEEE Standard 142-1991, IEEE Recommended Practice for Grounding of Industrial and Commercial Power Systems.

Kinyon, A. L.: Earth Resistivity Measurements for Grounding Grids, Trans. *AIEE, Power Appar. Syst.,* December 1961, vol. 80, pp. 795–800.

17.2 GAS-INSULATED SUBSTATIONS

By Philip Bolin

17.2.1 Introduction

High-voltage gas-insulated substations have been in service since the early 1960s. Operation of 800-kV equipment has proved successful since the end of 1979. Prototype testing of 1100 through 1600-kV substation equipment proved the feasibility of this equipment at the next generation of voltage levels.

17.2.2 General Characteristics

The basic principle of gas-insulated equipment is that the high-voltage current-carrying parts are within a metal enclosure and are held in a concentric configuration by cast epoxy spacer insulators. The space between the conductor and the enclosure is filled with sulfur hexafluoride gas under moderate pressure.

Medium-voltage to 170-kV equipment is available in three phases in one enclosure; for higher voltages, it is generally in a single-phase enclosure arrangement. The equipment can be installed indoors or outdoors, and it can be designed for any bus scheme. Depending on the voltage level, bus scheme, and whether connecting lines are installed underground or overhead, the land area required for gas-insulated equipment is 10% for 800 kV to 20% for 145 kV of the space required for comparable air-insulated equipment. Because of its smaller size and enclosed current-carrying parts, this equipment is excellently suited for installation where real estate is at a premium, where the environmental constraints dictate a minimum of visual exposure, and where the continuity of service may be threatened by airborne contamination. Typical section and five-bay substation layouts are shown in Figs. 17-26 and 17-27.

The dielectric medium is the sulfur hexafluoride (SF_6) gas, which became commercially available in 1947. SF_6 has been used as an insulating medium in electronic devices, power apparatus, and HVDC converter stations. Its excellent properties make it ideally suited both as an insulating and as an arc-quenching agent. SF_6 gas is colorless, odorless, chemically inert, nontoxic, nonflammable, and noncorrosive. Its dielectric strength is greatly superior to that of air, and it is close to 100 times as effective as air in quenching an electric arc. These characteristics are illustrated in Figs. 17-28 and 17-29, respectively.

Pure SF_6 is heavier than air, which causes it to settle in low areas, thus diluting oxygen in air. It is therefore necessary to learn proper safety rules before entering any area where pockets of SF_6 could accumulate. Although the gas is self-restoring, during its exposure to an electric arc it will yield decomposition by-products. In the presence of moisture, which is especially the case in failed and ruptured equipment, these by-products will hydrolyze, and all resulting reaction products must be considered hazardous.

The level of gas pressure at which the equipment will operate to meet specified ratings is a function of the relationship between diameters of the conductor and the enclosure (the size of the gap), and the temperature at which the equipment will operate. At the higher pressures, the gas would liquefy at higher temperatures, as indicated in Fig. 17-30. At lower pressures, dielectric strength and arc-quenching qualities of the gas would be reduced. Therefore, the gas-insulated equipment operating pressure is usually between 0.35 and 0.52 MPa (50 and 75 lb/in^2, gage).

Environmental effects of SF_6 that might be released to the atmosphere from GIS have been thoroughly studied. SF_6 does not affect the earth's ozone layer, but it is a strong greenhouse gas. Relative to CO_2, it has a global warming potential of 23,400 due to its infrared absorption and emission characteristics and very long life in the atmosphere (half-life is projected to be 3200 years). Fortunately, the concentration of SF_6 in the atmosphere is very low, and with proper handling, leak checking, and recycling, the contribution of SF_6 to anthropogenic global warming due to its use in electrical equipment can be kept below 0.1%.

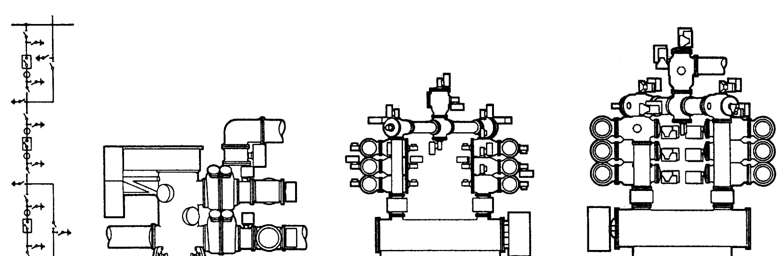


FIGURE 17-26 Typical breaker section for breaker-and-a-half scheme.

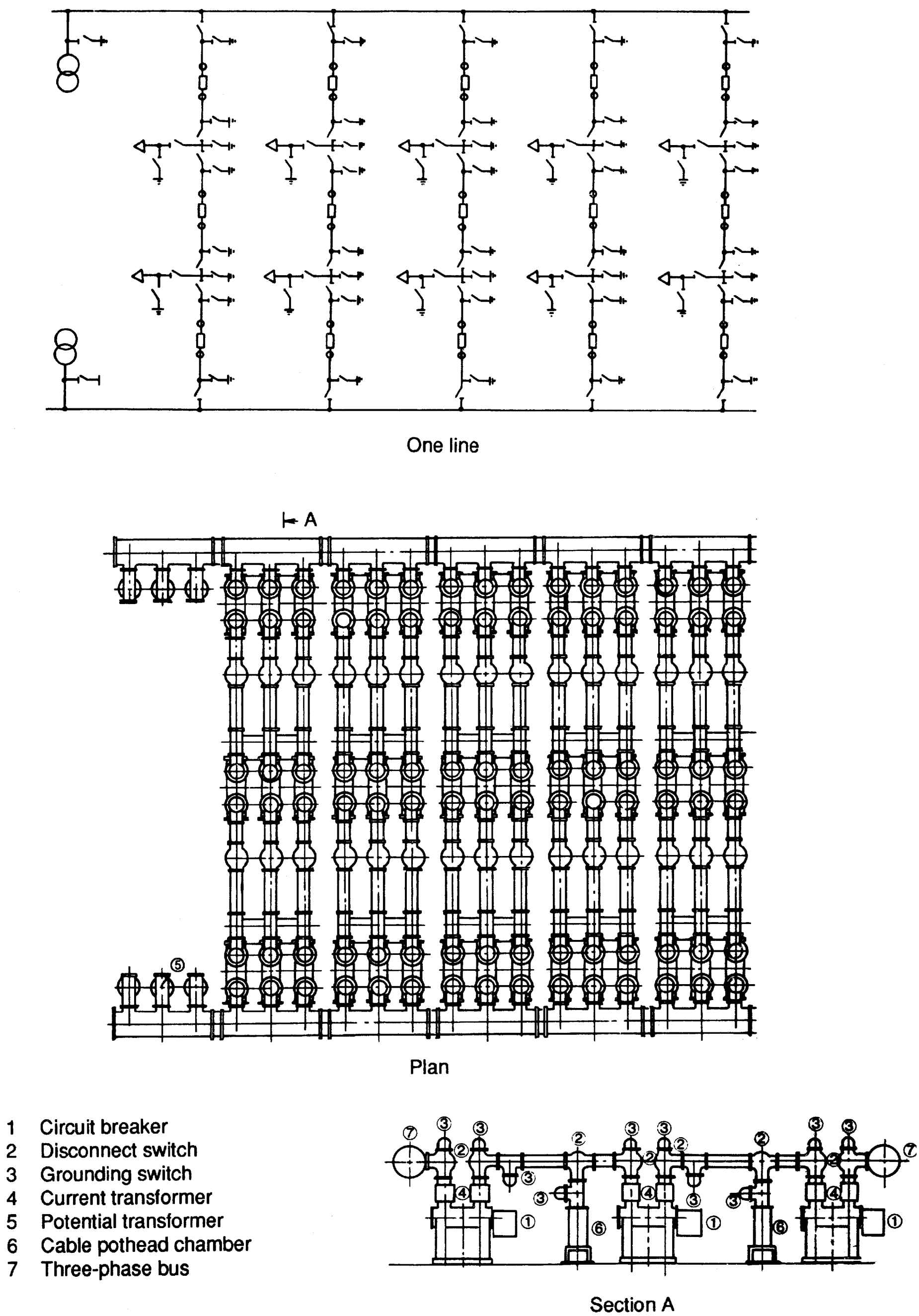


FIGURE 17-27 Typical layout for five-bay breaker-and-a-half scheme.

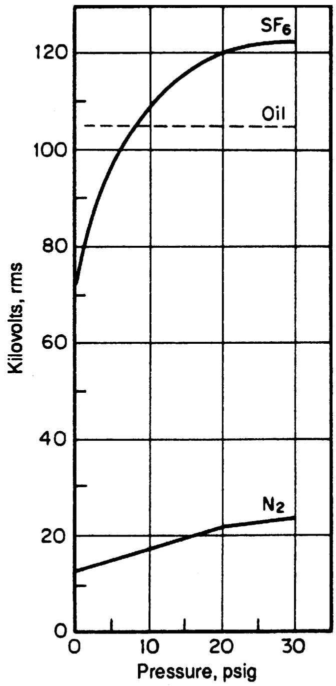


FIGURE 17-28 Power-frequency dielectric strength of SF_6.

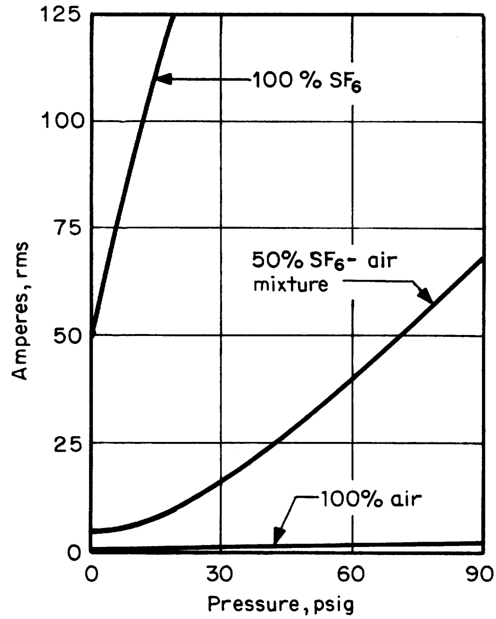


FIGURE 17-29 Arc-quenching ability of SF_6.

17.2.3 Equipment

The tubular conductor is made of aluminum or copper, and the enclosure can be of aluminum, steel, or stainless steel. The conductor connections are made by plug-in contacts, and the enclosure is joined by bolted flanges. In order to provide for proper gas seal, the flanges are constructed with O-ring gaskets. Conductor support insulators are of two types. Barrier insulators are used to isolate gas compartments; they must be capable of withstanding 1.5 times the operating pressure on one side and vacuum on the other side. Nonbarrier insulators permit the gas pressure to equalize between the compartments.

The circuit breakers are of dead-tank design and are the same as those installed in air-insulated substations, except that they are connected to the gas-insulated bus. They have standard ratings.

The rating of the disconnecting switches is established by standards; they must be capable of interrupting associated bus charging current. External indicators provide for switchblade position; however, visual verification is required by some users. This is done through a viewport in the enclosure directly over the contact-making area.

Maintenance and fault-closing grounding switches are the two most common types of grounding devices used with gas-insulated equipment. The first is used to provide grounding connection for maintenance purposes and is generally manually operated. The fault-closing grounding switch, in addition to providing for the maintenance function, has the capability of closing into a fault at least twice without damage. It is generally motor-operated. Both types may be furnished with a low-voltage test provision which permits voltage application to the conductor. This can be achieved without removal of the dielectric and without disassembly, except for ground shunt straps, which must be disconnected. Contact position indication can be the same as for disconnecting switches.

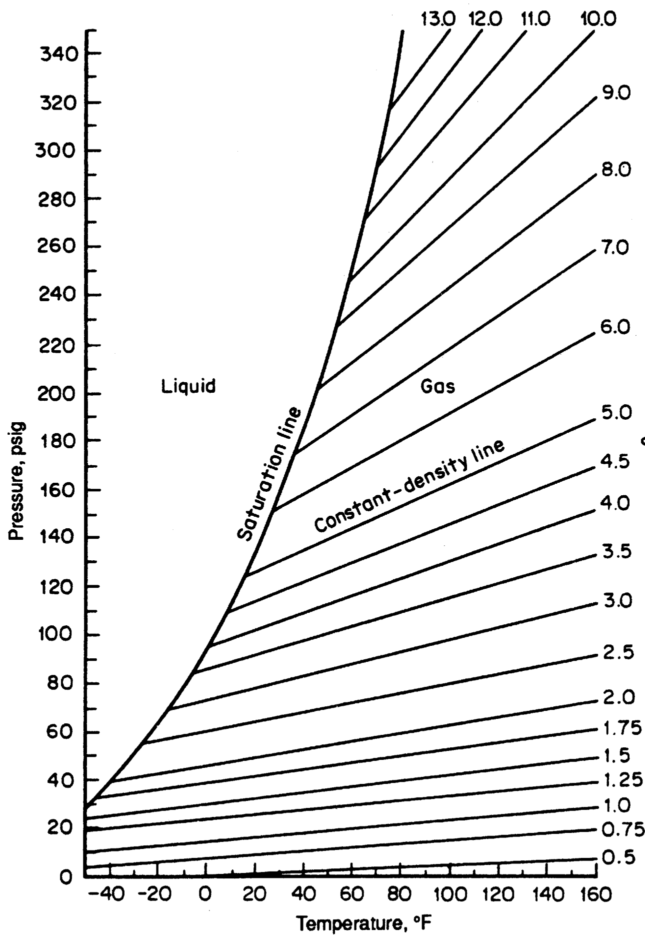


FIGURE 17-30 Pressure variation of SF_6 at constant specific volume.

Current transformers should be positioned so that the current in the enclosure does not affect the accuracy and ratio of the device and does not distort the conductor current being measured. Inductive and capacitive voltage transformers and surge arresters must be provided with disconnecting means for system dielectric tests. The surge-arrester ground connection must be isolated from the enclosure in order to permit monitoring of the leakage current.

The connections from the gas-insulated equipment to transmission lines, transformers, and reactors can be made by SF_6-to-air or SF_6-to-oil bushings. For overhead connection, the SF_6-to-air bushing is usually a hollow porcelain or composite insulator filled with pressurized SF_6 gas. There are two types of SF_6-to-oil bushings: one is for the transformer and the reactor and consists of an expanded bus section that totally encloses the bushing. Provision is made, for both the conductor and the enclosure, to minimize the transfer of transformer or reactor vibrations. The other type of SF_6-to-oil bushing is the power-cable pothead termination into the bus. This bushing must allow for power-cable disconnection from the gas-insulated bus to permit cable dc field testing. Both SF_6-to-oil bushings are provided with barriers which prevent oil migration into the switchgear. Bushings are also available for termination of a solid-dielectric cable into the GIS.

Particle Traps. Any loose conductive particles left within the enclosure can, when the equipment is energized, produce a flashover. When voltage is applied, these particles are moved by the alternating field in a random mode along the lower part of the enclosure. Eventually they reach a particle trap and through its slots fall into the zero-field region where they become permanently trapped and are rendered harmless. Particle traps are placed at the support insulators.

Pressure-Relief Devices. The enclosure is designed so that overpressure caused by internal faults is limited by pressure-relief devices. The location of these devices is such that when activated, the escaping ionized gases do not pose a hazard to personnel. The bursting pressure is coordinated with the rated gas pressure and the pressure rise caused by arcing. If the enclosure material or volume is sufficient to withstand expected overpressure during an internal fault, the pressure-relief devices may not be required.

Desiccant. Depending on the composition of the metal, insulators, and other materials within the equipment and expected moisture content of the dielectric, desiccant may be placed in selected locations to maintain the total moisture content at acceptable levels. It may be contained in especially designed canisters or may be built into spaces of the equipment. The desiccant is also useful in absorbing arcing-related gas by-products.

Expansion. Expansion joints provide for installation alignment and compensate for thermal expansion. If they are to facilitate alignment, they are locked in place when alignment is completed. If they are to compensate for thermal expansion, they are to have means to preserve mechanical integrity of the enclosure and the conductor.

Gas System. For maintenance and monitoring and to restrict damage and contamination in case of a fault, the gas system is divided by means of gas-barrier insulators into basic compartments: each circuit breaker, each terminal compartment, and each main bus section. Each gas compartment has a monitoring system for gas-density with two sets of contacts. Electrically independent contacts operate in two stages: an alarm to refill the gas, normally 5% to 10% below normal, and an alarm to indicate that the pressure has reached minimum level to support equipment ratings. By weight, the individual compartments are not to experience more than 1% leakage per year. The compartments are connected with external gas piping. The piping, which is made of corrosion-resistant material, must be isolated to prevent circulating currents. At each compartment, provisions are made for connecting moisture measurement instrumentation and the gas service cart.

Access. To facilitate maintenance, handholes or manholes, depending on the equipment size, are provided in the enclosure at locations where maintenance-prone devices are located. These gastight accesses are entered only after the dielectric has been evacuated and the compartment thoroughly ventilated.

Associated Systems. Most of the protective and control practices for air-insulated substations apply also for gas-insulated equipment. The principal difference is the requirement for online gas-density alarms for the gas-insulated substations. Another significant difference is that circuit-breaker reclosing is blocked for faults detected anywhere within gas-insulated equipment and its associated gas-insulated transmission-line exits.

In considering the grounding of the gas-insulated equipment, it is essential that the enclosure be bonded so as to present a continuous current path. The current in the conductor induces a voltage in its single-phase enclosure, which causes a longitudinal current flow. When the loads are balanced, this current returns through the enclosure of the adjacent phase. A discontinuity in the enclosure would generate circulating currents and most likely higher-than-desired touch potential.

Field testing of gas-insulated substations requires tests, which may not be required for the conventional equipment. These tests are leak detection, moisture content in the dielectric, and power-frequency testing. In addition to verifying the integrity of the installation, the power-high-voltage frequency test also will reveal the presence of any free conducting particles which may be present.

REFERENCES ON SF_6 GAS-INSULATED SUBSTATIONS

Bennett, A. I., Carlson, G. L., and Lee, A.: Characterization of ARC By-Products of Sulfur Hexafluoride and Polymeric Construction Materials, *EPRI Final Report* EL-5089, April 1987.

Boggs, S., Chu, F. Y., and Fujimoto, J., ed.: *Gas-Insulated Substations—Technology and Practices.* New York: Pergamon Press, 1986.

CIGRE Working Group 23-02 Report: Handling of SF_6 and Its Decomposition Products in Gas-Insulated Switchgear, CIGRE, *Electra,* no. 136, June 1991, part 1, and no. 137, August 1991, part 2.

CIGRE Working Group 23.10: A Twenty-Five-Year Review of Experience with SF_6 Gas-Insulated Substations (GIS), CIGRE Paper 23.101, 1992.

CIGRE Working Group 23.10-01 Report 117: SF_6 Recycling Guide, August 1997.

IEEE C37.122-1993, IEEE Standard for Gas Insulated Substations.

IEEE C37.123-1996, IEEE Guide to Specifications for Gas-Insulated Substations Equipment.

IEEE C37.38-1989, IEEE Standard for Gas-Insulated, Metal-Enclosed Disconnecting, Interrupter, and Grounding Switches.

SECTION 18
POWER DISTRIBUTION

Daniel J. Ward
Principal Engineer, Dominion Virginia Power; Fellow, IEEE; Chair, IEEE Distribution Subcommittee; Chair, ANSI C84.1 Committee, Past Vice Chair (PES), Power Quality Standards Coordinating Committee

CONTENTS

18.1 DISTRIBUTION DEFINED

Broadly speaking, *distribution* includes all parts of an electric utility system between bulk power sources and the consumers' service-entrance equipments. Some electric utility distribution engineers, however, use a more limited definition of distribution as that portion of the utility system between the distribution substations and the consumers' service-entrance equipment. In general, a typical distribution system consists of (1) subtransmission circuits with voltage ratings usually between 12.47 and 345 kV which deliver energy to the distribution substations, (2) distribution substations which convert the energy to a lower *primary system* voltage for local distribution and usually include facilities for voltage regulation of the primary voltage, (3) primary circuits or *feeders*, usually operating in the range of 4.16 to 34.5 kV and supplying the load in a well-defined geographic area, (4) distribution transformers in ratings from 10 to 2500 kVA which may be installed on poles or grade-level pads or in underground vaults near the consumers and transform the primary voltages to utilization voltages, (5) secondary circuits at utilization voltage which carry the energy from the distribution transformer along the street or rear-lot lines, and (6) service drops which deliver the energy from the secondary to the user's service-entrance equipment. Figures 18-1 and 18-2 depict the component parts of a typical distribution system.

Distribution investment constitutes 50% of the capital investment of a typical electric utility system. Recent trends away from generation expansion at many utilities have put increased emphasis on distribution system development.

The function of distribution is to receive electric power from large, bulk sources and to distribute it to consumers at voltage levels and with degrees of reliability that are appropriate to the various types of users.

For single-phase residential users, American National Standard Institute (ANSI) C84.1-1989 defines *Voltage Range A* as 114/228 V to 126/252 V at the user's service entrance and 110/220 V to 126/252 V at the point of utilization. This allows for voltage drop in the consumer's system. Nominal voltage is 120/240 V. Within Range A utilization voltage, utilization equipment is designed and rated to give fully satisfactory performance.

As a practical matter, voltages above and below Range A do occur occasionally; however, ANSI C84.1 specifies that these conditions shall be limited in extent, frequency, and duration. When they do occur, corrective measures shall be undertaken within a reasonable time to improve voltages to meet Range A requirements.

Rapid dips in voltage which cause incandescent-lamp "flicker" should be limited to 4% or 6% when they occur infrequently and 3% or 4% when they occur several times per hour. Frequent dips, such as those caused by elevators and industrial equipment, should be limited to 1½% or 2%.

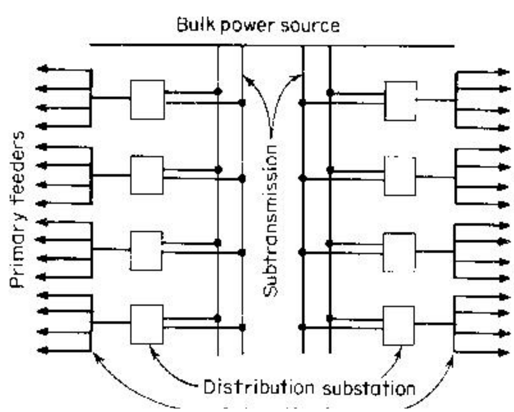


FIGURE 18-1 Typical distribution system.

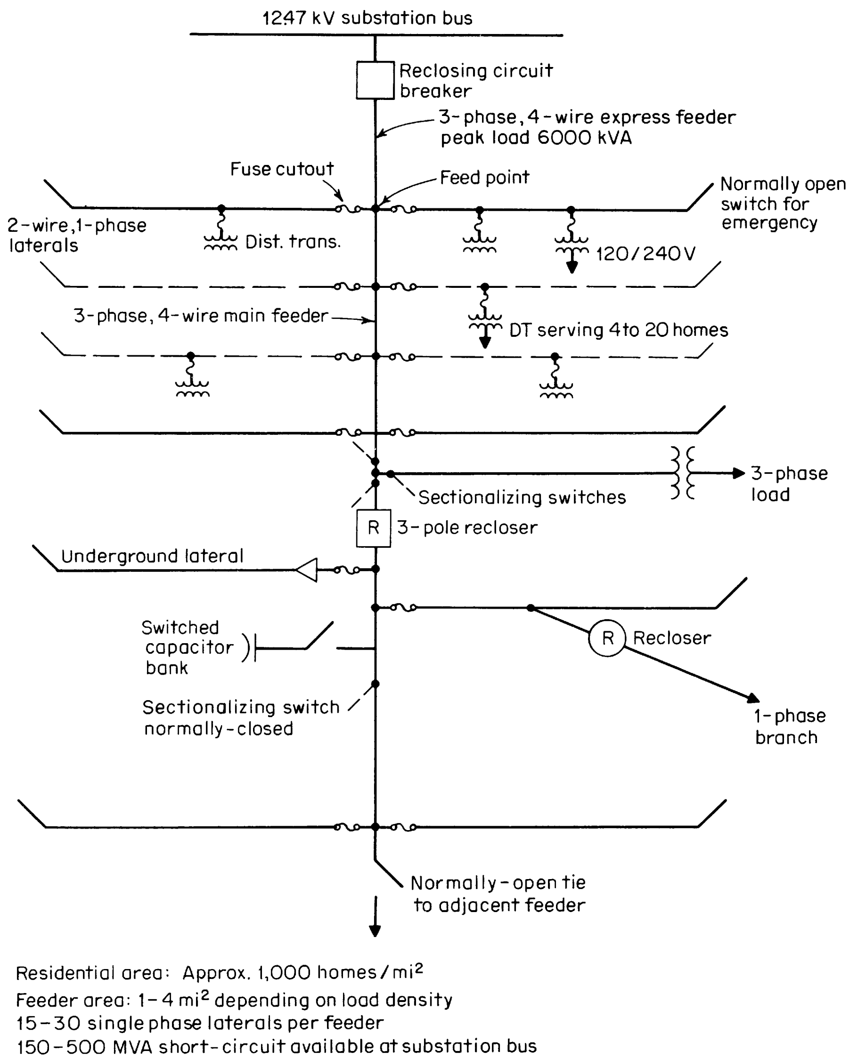


FIGURE 18-2 One-line diagram of typical primary distribution feeder.

Reliability of service can be described by factors such as frequency and duration of service interruptions. While short and infrequent interruptions may be tolerated by residential and small commercial users, even a short interruption can be costly in the case of many industrial processes and can be dangerous in the case of hospitals and public buildings. For such sensitive loads, special measures are often taken to ensure an especially high level of reliability, such as redundancy in supply circuits and/or supply equipment. Certain computer loads may be sensitive not only to interruptions but even to severe voltage dips and may require special power-supply systems which are virtually uninterruptible.

From a system-planning and design point of view, the optimal choice of subtransmission voltage and system arrangement is closely interrelated with distribution substation size and with the primary distribution voltage level. At any given time, the most economical arrangement is achieved when the sum of the subtransmission, substation, and primary feeder costs to serve an area is a minimum over the life of the facilities. In practice, the number, size, and availability of bulk supply sources for feeding the subtransmission may be significant factors as well.

A distribution system should be designed so that anticipated load growth can be served at minimum expense. This flexibility is needed to handle load growth in existing areas as well as load growth in new areas of development.

Overhead and underground distribution systems are both used in large metropolitan areas. In the past in smaller towns and in the less-congested areas of larger cities, overhead distribution was almost universally used; the cost of underground distribution for residential areas was several times that of overhead. During the past 25 to 30 years, the cost of underground residential distribution (URD) has been reduced drastically through the development of low-cost, solid-dielectric cables suitable for direct burial, mass production of pad-mounted distribution transformers and accessories, mechanized cable-installation methods, etc. The cost of a typical URD system for a new residential subdivision is about 50% greater than that of an overhead system in many areas; in others, there is little or no differential due to local land conditions. As a result, some utilities will justifiably have some type of extra charge for underground. With the increased public interest in improving the appearance of residential areas and the declining cost of URD, the growth of URD has been extremely rapid. Today, perhaps as much as 70% of new residential construction is served underground. A number of states have enacted legislation making underground distribution mandatory for new residential subdivisions.

***Undergrounding*[*].** In the last decade, U.S. East Coast and Midwest regions experienced several catastrophic "100 year storms." These storms left widespread electric power outages that lasted several days. Given the critical role that electricity plays in our modern lifestyle, even a momentary power outage is an inconvenience. A days-long power outage presents a major hardship and can be catastrophic in terms of its health and safety consequences, and the economic losses it creates. Why then, don't we bury more of our power lines so they will be protected from storms?

The fact is we already are placing significant numbers of power lines underground. Over the past 10 years, approximately half of the capital expenditures by U.S. investor-owned utilities for new transmission and distribution wires have been for underground wires. Almost 80% of the nation's electric grid, however, has been built with overhead power lines. Would electric reliability be improved if more of these existing overhead lines were placed underground as well?

What the report finds is that burying existing overhead power lines does not completely protect consumers from storm-related power outages. However, underground power lines do result in fewer overall power outages, but the duration of power outages on underground systems tends to be longer than for overhead lines. Also, undergrounding is expensive, costing up to $1 million/mile or almost 10 times the cost of a new overhead power line. This means that most undergrounding projects cannot be economically justified and must cite intangible, unquantifiable benefits such as improved community or neighborhood aesthetics for their justification. Determining who pays and who benefits from undergrounding projects can be difficult and often requires the establishment of separate government-sponsored programs for funding.

How Much Does Undergrounding Improve Electric Reliability? Comparative reliability data indicate that the frequency of outages on underground systems can be substantially less than for overhead systems. However, when the duration of outages is compared, underground systems lose much of their advantage. The data show that the frequency of power outages on underground systems is only about one-third of that of overhead systems. A 2000 report issued by the Maryland Public Service Commission concluded that the impact of undergrounding on reliability was "unclear."

In a 2003 study, the North Carolina Commission summarized 5 years of underground and overhead reliability comparisons for North Carolina's investor-owned electric utilities–Dominion North Carolina Power, Duke Energy, and Progress Energy Carolinas. The data indicate that the frequency of outages on underground systems was 50% less than for overhead systems, but the average duration of an underground outage was 58% longer than for an overhead outage. In other words, for

[*]From "Out of Sight, Out of Mind?," January 2004, Edison Electric Institute (used with permission).

the North Carolina utilities, an underground system suffers only about half the number of outages of an overhead system, but those outages take 1.6 times longer to repair. Based on this data, Duke Power concluded, "Underground distribution lines will improve the potential for reduced outage interruption during normal weather, and limit the extent of damage to the electrical distribution system from severe weather-related storms." However, once an interruption has occurred, underground outages normally take significantly longer to repair than a similar overhead outage.

Reliability Characteristics of Overhead and Underground Power Lines

- Overhead lines tend to have more power outages primarily due to trees coming in contact with overhead lines.
- It is relatively easy to locate a fault on an overhead line and repair it. A single line worker, for example, can locate and replace a blown fuse. This results in shorter duration outages.
- Underground lines require specialized equipment and crews to locate a fault, a separate crew with heavy equipment to dig up a line, and a specialized crew to repair the fault. This greatly increases the cost and the time to repair a fault on an underground system.
- In urban areas, underground lines are 4 times more costly to maintain than overhead facilities.
- Underground lines have a higher failure rate initially due to dig-ins and installation problems. After 3 or 4 years, however, events that affect failures become virtually nonexistent.
- As underground cables approach their end of life, failure rates increase significantly and these failures are extremely difficult to locate and repair. Maryland utilities report that their underground cables are becoming unreliable after 15 to 20 years and reaching their end of life after 25 to 35 years.
- Pepco found that customers served by 40-year-old overhead lines had better reliability than customers served by 20-year-old underground lines.
- Two Maryland utilities have replaced underground distribution systems with overhead systems to improve reliability.
- Water and moisture infiltration can cause significant failures in underground systems when they are flooded, as often happens in hurricanes.
- Due to cost or technical considerations, it is unlikely that 100% of the circuit from the substation to the customer can be placed entirely underground. This leaves the circuit vulnerable to the same types of events that impact other overhead lines, for example, high winds and ice storms.

Other Benefits of Undergrounding. One of the most commonly cited benefits of undergrounding is the removal of unsightly poles and wires. Local communities and neighborhoods routinely spend millions to place their existing overhead power lines underground.

Similarly, when given the option, builders of new residential communities will often pay a premium of several thousand dollars/home to place the utilities underground. These "aesthetic" benefits are virtually impossible to quantify, but are, in many instances, the primary justification for projects to place existing power lines underground.

Underground lines do have other benefits. In 1998, Australia completed a major benefit/cost analysis of undergrounding all existing power lines in urban and suburban areas throughout the country. The study costed more than $1.5 million Australian ($1.05 million U.S. at current rates), and represents what may be the most comprehensive undertaking to date to quantify the benefits and costs related to undergrounding.

In addition to the value of improved aesthetics, the study identified the following potential benefits related to undergrounding that it attempted to quantify:

- Reduced motor vehicle accidents caused by collisions with poles
- Reduced losses caused by electricity outages
- Reduced network maintenance costs
- Reduced tree-pruning costs

- Increased property values
- Reduced transmission losses due to the use of larger conductors
- Reduced greenhouse-gas emissions (lower transmission losses)
- Reduced electrocutions
- Reduced brushfire risks, and
- Indirect effects on the economy such as employment

Of this list, the only four items deemed significant in the study's benefit/cost calculations included:

- Motor vehicle accidents
- Maintenance costs
- Tree-trimming costs, and
- Line losses

The Australian list of benefits does not include improved reliability as a significant benefit of undergrounding. Instead it identifies the reduction in losses from motor vehicle accidents as the largest benefit from undergrounding—something utilities have no control over.

Underground cost data for U.S. utilities indicate that the cost of placing overhead power lines underground is 5 to 10 times the cost of new overhead power lines. Other factors also can result in substantial additional customer costs for undergrounding projects. These include:

- Electric undergrounding strands other utilities, for example, cable and telephone companies, which must assume 100% of pole costs if electric lines are underground. These additional nonelectric costs will likely be passed on to cable and telephone consumers.
- Customers may incur substantial additional costs to connect homes to newly installed underground service, possibly as much as $2000 if the household electric service must be upgraded to conform to current electric codes.

Paying for Undergrounding. In spite of its high cost and lack of economic justification, undergrounding is very popular across the country. In 9 out of 10 new subdivisions, contractors bury power lines. In addition, dozens of cities have developed comprehensive plans to bury or relocate utility lines to improve aesthetics.

For new residential construction, utilities vary on how they charge for the cost of providing underground services. When it comes to converting existing overhead lines to underground, a variety of programs are being utilized. They include special assessment areas, undergrounding districts, and state and local government initiatives.

Placing existing power lines underground is expensive, costing approximately $1 million/mile. This is almost 10 times the cost of a new overhead power line.

While communities and individuals continue to push for undergrounding—particularly after extended power outages caused by major storms—the reliability benefits that would result are uncertain, and there appears to be little economic justification for paying the required premiums.

Indeed, in its study of the undergrounding issue, the Maryland Public Service Commission concluded, "If a 10 percent return is imputed to the great amounts of capital freed up by building overhead instead of underground lines, the earnings alone will pay for substantial ongoing overhead maintenance," implying that utilities could have more resources available to them to perform maintenance and improve reliability on overhead lines if they invested less in new underground facilities.

For the foreseeable future, however, it appears that the undergrounding of existing overhead power lines will continue, justified primarily by aesthetic considerations—not reliability or economic benefits. Many consumers simply want their power lines placed underground, regardless of the costs. The challenge for decision makers is determining who will pay for these projects and who will benefit.

There are several undergrounding programs around the country that are working through these equity issues and coming up with what appear to be viable compromises. Once a public-policy decision is reached to pursue an undergrounding project, it is worthwhile for the leaders involved to evaluate these programs in more detail to determine what is working, and what is not.

Rural Service. Rural service has been extended to most farmers and rural dwellers through the efforts of utilities, cooperatives, and government agencies. Rural construction must be of the least-expensive type consistent with durability and reliability because there may be only a few users per mile of line. Historically, rural construction has been overhead, but the advent of cable-plowing techniques has made underground economically competitive with overhead in some parts of the country, and a growing amount of rural distribution is being installed underground.

Higher primary voltages of 24.9Y/14.4 and 34.5Y/19.92 kV are continuing to grow in usage, although primary voltages in the 15-kV class predominate. The 5-kV class continues to decline in usage. Surveys indicate that in recent years approximately 78% of the overhead and underground line additions are at 15 kV, 11% are at 25 kV, and 7.5% are at 35 kV.

Generally, when a higher distribution voltage is initiated, it is built in new, rapidly growing load areas. The economic advantage of the higher voltages usually is not great enough to justify massive conversions of existing lower-voltage facilities to the higher level. The lower-voltage areas are contained and gradually compressed over a period of years as determined by economics, obsolescence, and convenience. Virtually, all modern primary systems serving residential and small commercial and small industrial loads are 4-wire, multigrounded, common-neutral systems.

18.2 DISTRIBUTION-SYSTEM AUTOMATION

Distribution automation (DA), a system to monitor and control the distribution system in real-time, was gradually introduced in the 1970s more as a concept than a fully developed plan. Unlike the introduction of EMS, where utilities readily saw the benefits of automatic generation control and economic dispatch and adopted the technology, utilities were much more cautious in their approach to distribution automation.

Early distribution automation projects were undertaken by a handful of utilities. The technology was changing and evolving so much so that DA was being touted as an amorphous system capable of covering any imaginable function under the sun. A 1984 EPRI project, Guidelines for Evaluating Distribution Automation, focused attention on what functions could be automated and what value could be attached to those functions. A positive result of this project is that it got people thinking about what functions mattered most. However, it was a little bit ahead of its time in that there wasn't much standardization in systems employed for DA and one couldn't simply select functions of interest and expect to obtain a system that could be built for the total value of the functions selected. Then too, the choice of the communications systems (e.g., telephone, fiber optics, radio, carrier, etc.) proved to be a barrier to widespread implementation.

At the substation level, equipment loadings became an early focus, and asset management became a desired function for DA systems. In addition, the ability to trip distribution circuit breakers and transfer load between substations was commonplace as SCADA was added and this represented the extent of distribution automation to many companies.

Volt/var control, that is, controlling the combination of load tap changers (LTC) or voltage regulators and switched capacitor banks within a substation, was a function many companies incorporated with DA. With adoption of microprocessor relays and fault distance relaying, some incorporated the output information from fault distance relays and diagnostic alarms from various subsystems to be part of the DA package.

Moving outside the substation, controlling automated circuit tie switches was prompted by reliability considerations. Having SCADA links to other reclosers, particularly the ones with microprocessor controls, enabled more ability to remotely control field switching and achieve more rapid restoration of service.

Distribution automation is still evolving with systems incorporating many of the functions previously described. More utilities are employing varying degrees of distribution automation and more standardization is taking place.

18.3 CLASSIFICATION AND APPLICATION OF DISTRIBUTION SYSTEMS

Distribution systems may be classified in according to:

- voltage—120 V, 12,470 V, 34,500 V, etc.
- scheme of connection—radial, loop, network, multiple, and series.
- loads—residential, small light and power, large light and power, street lighting, railways, etc.
- number of conductors—2-wire, 3-wire, 4-wire, etc.
- type of construction—overhead or underground.
- number of phases—single-phase, 2-phase, or 3-phase; and as to frequency: 25 Hz, 60 Hz, etc.

Application of Systems. In American practice, alternating-current (ac) 60-Hz systems are almost universally used for electric power distribution. These systems comprise the most economical method of power distribution, owing in large measure to the ease of transforming voltages to levels appropriate to the various parts of the system. These transformations are accomplished by means of reliable and economical transformers. By proper system design and the application of overvoltage and overcurrent protective equipment, voltage levels and service reliability can be matched to almost any consumer requirement.

Single-phase residential loads generally are supplied by simple radial systems at 120/240 V. The ultimate in service reliability is provided in densely loaded business/commercial areas by means of grid-type secondary-network systems at 208Y/120 V or by "spot" networks, usually at 480Y/277 V. Secondary-network systems are used in about 90% of the cities in this country having a population of 100,000 or more and in more than one-third of all cities with populations between 25,000 and 100,000.

Where secondary-network systems do not supply sufficiently reliable service for critical loads, emergency generators and/or batteries are sometimes provided together with automatic switching equipment so that service can be maintained to the critical loads in the event that the normal utility supply is interrupted. Such loads are found in hospitals, computer centers, key industrial processes, etc.

Single-phase residential loads are almost universally supplied through 120/240-V, 3-wire, single-phase services. Large appliances, such as ranges, water heaters, and clothes dryers, are served at 240 V. Lighting, small appliances, and convenience outlets are supplied at 120 V.

An exception to the preceding comments occurs when the dwelling unit is in a distributed secondary-network area served at 280Y/120 V. In this case, large appliances are supplied at 208 V and small appliances at 120 V.

Three-phase, 4-wire, multigrounded, common-neutral primary systems, such as 12.47Y/7.2 kV, 24.9Y/14.4 kV, and 34.5Y/19.92 kV, are used almost exclusively. The fourth wire of these Y-connected systems is the neutral for both the primary and the secondary systems. It is grounded at many locations. Single-phase loads are served by distribution transformers, the primary windings of which are connected between a phase conductor and the neutral. Three-phase loads can be supplied by 3-phase distribution transformers or by single-phase transformers connected to form a 3-phase bank. Primary systems in the 15-kV class are most commonly used, but the higher voltages are gaining acceptance. Figure 18-2 illustrates a typical radial primary feeder.

The 4-wire system is particularly economic for URD systems because each primary lateral or branch circuit consists of only one insulated phase conductor and the bare, uninsulated neutral rather than two insulated conductors. Also, only one primary fuse is required at each transformer and one surge arrester in overhead installations.

Three-phase, 3-wire primary systems are not widely used for public distribution, except in California. They can be used to supply single-phase loads by means of distribution transformers having primary winding connected between two phase conductors. Single-phase primary laterals consist of two insulated phase conductors; each single-phase distribution transformer requires two fuses and two surge arresters (where used). Three-phase loads are served through 3-phase distribution transformers or appropriate 3-phase banks. Two-phase systems are rarely used today.

18.4 CALCULATION OF VOLTAGE REGULATION AND I^2R LOSS

When a circuit supplies current to a load, it experiences a drop in voltage and a dissipation of energy in the form of heat. In dc circuits, voltage drop is equal to current in amperes multiplied by the resistance of the conductors, $V = IR$. In ac circuits, voltage drop is a function of load current and power factor and the resistance and reactance of the conductors. Heating is caused by conductor losses; for both dc and ac circuits they are computed as the square of current multiplied by conductor resistance in ohms. Watts $= I^2R$, or kW $= I^2R/1000$. Capacitance can usually be neglected for calculation in distribution circuits because its effect on voltage drop is negligible for the circuit lengths and operating voltages used. In circuit design, a conductor size should be selected so that it will carry the required load within specified voltage-drop limits and will have an optimized value of installed cost and cost of losses. Today, a conductor size meeting these criteria will operate well within safe operating temperature limits. In some cases, short-circuit current requirements will dictate the minimum conductor size.

Percent voltage drop or percent regulation is the ratio of voltage drop in a circuit to voltage delivered by the circuit, multiplied by 100 to convert to percent. For example, if the drop between a transformer and the last customer is 10 V and the voltage delivered to the customer is 240, the percent voltage drop is $10/240 \times 100 = 4.17\%$. Often the nominal or rated voltage is used as the denominator because the exact value of delivered voltage is seldom known.

Percent I^2R or percent conductor loss of a circuit is the ratio of the circuit I^2R or conductor loss, in kilowatts, to the kilowatts delivered by the circuit (multiplied by 100 to convert to percent). For example, assume a 240-V single-phase circuit consisting of 1000 ft of two No. 4/0 copper cables supplies a load of 100 A at unity power factor.

$$I^2R = 100^2 \times 2 \times 0.0512 = 1024 \text{ W} = 1.024 \text{ kW}$$

$$\text{Load delivered} = 240 \times 100 = 24{,}000 \text{ W} = 24 \text{ kW}$$

$$\%\ I^2R \text{ loss} = 1.024/24 \times 100 = 4.26\%$$

Direct-current voltage drop is easily calculated by multiplying load amperes I by ohmic resistance R of the conductors through which the current flows (see Sec. 4 for ohmic resistance of various conductors).

Example: A 500-ft dc circuit of two 4/0 copper cables carries 200 A. What is the voltage drop? Resistance of 1000 ft of 4/0 copper cable is 0.0512 Ω.

$$\text{Drop} = IR = 200 \times 0.0512 = 10.24 \text{ V}$$

If 240 is the delivered voltage,

$$\%\ \text{regulation} = 10.24/240 \times 100 = 4.26\%$$

I^2R or conductor loss in dc or ac circuits is calculated by multiplying the square of the current in amperes by ohmic resistance of the conductors through which the current flows. The result is in watts.

In dc circuits, percent voltage drop and percent conductor loss are identical.

$$\%\ \text{voltage drop} = IR/V \times 100$$

$$\%\ I^2R = I^2R/VI \times 100 = IR/V \times 100$$

In ac circuits, the ratio of percent conductor loss to percent voltage regulation is given approximately by the following approximate formula:

$$\frac{\%\ I^2R \text{ loss}}{\%\ \text{voltage drop}} = \frac{\cos\phi}{\cos\theta\,\cos(\phi - \theta)} \tag{18-1}$$

where θ = power-factor angle and ϕ = impedance angle; that is, $\tan\phi = X/R$.

TABLE 18-1 Voltage Drop in Volts per 100,000 A · ft, 2-Wire DC Circuits (Loop)

Conductor size, AWG or kcmil		Volts drop per 100,000 A · ft, 90º copper temp
Copper	Approx. equivalent aluminum	
6	4	102.8
4	2	64.6
2	1/0	40.7
1/0	3/0	25.6
2/0	4/0	20.3
4/0	336	12.8
350	556	7.71
500	795	5.39
1000		2.70
1500		1.80
2000		1.35

Note: 1 ft = 0.3048 m.

Table 18-1 gives voltage drop in volts per 100,000 A · ft for 2-wire dc circuits for a number of conductor sizes. Ampere-feet is the product of the number of amperes of current flowing and the distance in feet between the sending and receiving terminals multiplied by 2 to take into account the drop in both the outgoing and return conductors. Or the feet can be considered to be the total number of conductor feet, outgoing and return.

Table 18-1 also gives the voltage drop for 3-wire circuits when serving balanced loads, where the term "feet" is taken to mean twice the number of feet between sending and receiving terminals.

Example 1. What is the voltage drop and percent voltage drop when 200 A dc flows 1500 ft one way through a 2-wire, 120-V, 556-kcmil aluminum circuit? First determine ampere-feet factor as 100 × 1500/100,000 = 1.5. From Table 18-1, the voltage drop is 7.71 V per 100,000 A · ft. This value multiplied by the 1.5 factor gives the total voltage drop = 1.5 × 7.71 = 11.6 V. The percent voltage drop = 11.6 × 100/120 = 9.64%. The percent conductor loss also is 9.64%, which is equivalent to 120 × 100 × 0.0954 = 1.16 kW.

Example 2. A mine 1 mile from a motor-generator station must receive 100 kW dc at not less than 575 V. Maximum voltage of the generator is 600 V. What conductor size should be used?

$$\text{Max. current} = \frac{100{,}000\ \text{W}}{575\ \text{V}} = 173.9\ \text{A}$$

$$\text{Loop ft} = 2 \times 5280 = 10{,}560\ \text{ft}$$

$$\frac{\text{A} \cdot \text{ft}}{100{,}000} = \frac{173.9 \times 10{,}560}{100{,}000} = 18.36$$

$$18.36 \times \text{voltage drop per } 100{,}000\ \text{A} \cdot \text{ft from Table 18-1} = 25\ \text{V}$$

Therefore, voltage drop per 100,000 A · ft = 25/18.36 = 1.36. From Table 18-1, the copper conductor size corresponding to 1.36 V/100,000 A · ft is 2000 kcmil copper.

Calculating Voltage Drop in AC Circuits. The voltage drop per mile in each round wire of 3-phase 60-Hz line with equilateral spacing D inches between centers or in each wire of a single-phase line D inches between centers is

$$\tilde{V}\,\text{drop} = \tilde{I}R + j\tilde{I}\left(0.2794 \log \frac{D}{r} + 0.03034\,\mu\right) \quad \text{volts in phasor form} \qquad (18\text{-}2)$$

where $\tilde{I}$ is in phasor amperes, R is the 60-Hz resistance of the wire per mile, Ω, log is the log to base 10, r is the radius of round wire, in, and μ is the permeability of the wire (unity for nonmagnetic materials such as copper or aluminum). j in Eq. (18-2) denotes an angle of 90°; $+j$ means 90° leading, $-j$ means 90° lagging. Thus, the expression for phasor current lagging the reference voltage is $\tilde{I} = I_x - jI_y = I\sqrt{\theta^\circ}$ with reference to a conveniently chosen horizontal axis of reference—usually sending- or receiving-end voltage. The symbol ~ over I or V indicates phasor values. Voltage drops determined in this manner are also phasors and are with respect to the reference axis.

When wire is stranded, an equivalent radius must be used for r in Eq. (18-2). $r = 0.528\ \sqrt{A}$ for 7 strands, $r = 0.5585\ \sqrt{A}$ for 19 strands, $r = 0.5675\ \sqrt{A}$ for 37 strands, where r = equivalent radius, in, and A = area of metal, in^2.

Frequency is 60 Hz for the constants in parentheses in Eq. (18-2), which gives reactance X in ohms per mile. For 25 Hz, multiply by 25/60. The equation is sometimes written

$$\tilde{V} \text{ drop per mile} = \tilde{I}\,(R + jX) = \tilde{I}\tilde{Z} \qquad \text{volts in phasor form} \tag{18-3}$$

where I is in phasor amperes and $Z = Z/\phi \cdot \Omega$/mi at 60 Hz.

Three unsymmetrically spaced wires a, b, and c of a 3-phase circuit with correct transpositions can have voltage drop in each wire calculated by Eq. (18-2) by substituting for D the geometric mean of the three interaxial distances:

$$D = \sqrt[3]{D_{ab}D_{bc}D_{ca}}$$

The Phasor Method. In Eq. (18-3), I is in vector amperes,

$$\tilde{I} = I_x - jI_y = I\angle\theta$$

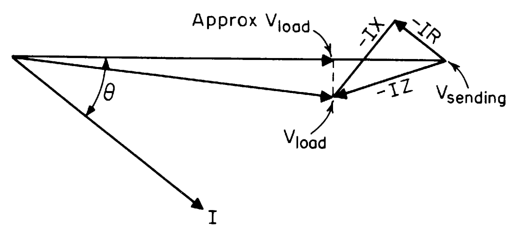


FIGURE 18-3 Phasor diagram showing voltage relationships.

where θ is the angle that the current lags (or leads) the voltage. The sending-end voltage is usually chosen as the axis, or phasor, of reference in drawing the phasor diagram. For example, consider Fig. 18-3, where sending voltage $V_s = V_s\underline{/\theta^\circ}$, load current $I = I\underline{/\theta^\circ}$, circuit impedance $\tilde{Z} = Z\underline{/\theta^\circ} = R + jX$, and load voltage $\tilde{V}_L = \tilde{V}_s - \tilde{I}\tilde{Z}$ (all phasors). The symbol $\angle$ is used for positive angles, assuming that the counterclockwise direction from the phasor or reference is positive and the clockwise directions negative. Assume that $V_s = 230/0^\circ$, $\tilde{I} = 50\underline{/-36.87^\circ}$, $\tilde{Z} = 0.2\underline{/71.57^\circ}$, and $\tilde{Z} = R + jX$. Thus

$$\begin{aligned}
\tilde{V}_L &= 230/0^\circ - 50\,\underline{/36.87^\circ} \times 0.2\underline{/71.57^\circ} \\
&= 230\underline{/\theta^\circ} - 10\underline{/34.70^\circ} \\
&= 230 - 10\cos 34.70^\circ - j\,10 \sin 34.70^\circ \\
&= 230 - 8.22 - j\,5.69 = 221.78 - j\,5.69 \\
&= 221.78 \text{ (very nearly)}
\end{aligned}$$

Neglecting the term $-j\,5.69$ simplifies the final calculation and gives the load voltage within a fraction of 1% of the precise result. This method is sufficiently accurate for practically all distribution engineering calculations and can be thought of as

$$V \text{ drop} = IR\cos\theta + IX\sin\theta = IZ\cos(\phi - \theta) \tag{18-4}$$

where I and Z are absolute magnitudes, not phasor quantities, ϕ is the impedance angle, and θ is the power-factor angle by which the current lags (or leads) the voltage. Calculating the drop in the above example by this method:

$$V \text{ drop} = 50 \times 0.2 \times \cos 71.57° \times \cos 36.87°$$
$$+ 50 \times 0.2 \times \sin 71.57° \times \sin 36.87°$$
$$= 2.53 + 5.69 = 8.22 \text{ V}$$

or

$$V \text{ drop} = IZ \cos(\phi - \theta) = 50 \times 0.2 \times \cos(71.57° - 36.87°)$$
$$= 10 \cos(34.7°) = 10 \times 0.822 = 8.22 \text{ V}$$

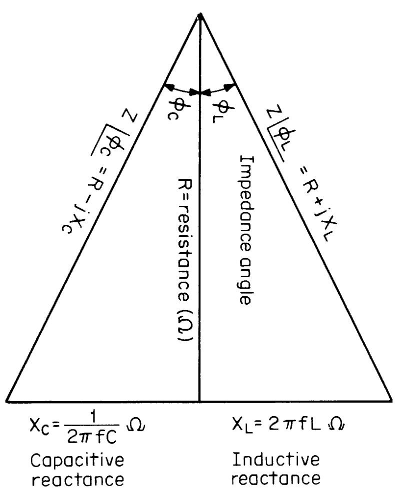


FIGURE 18-4 Impedance diagrams for series connection of resistance and reactance (L = inductance, in henrys; C = capacitance, in farads; F = frequency, in hertz).

Impedance Z can be visualized as the hypotenuse of a right triangle in which the base is the resistance R and the altitude is the reactance X. In phasor form, $\tilde{Z} = R \pm jX$, where the positive sign is used for inductive reactance and the negative sign for capacitive reactance. Impedance also can be expressed as $\tilde{Z} = Z\underline{/\phi°}$, where Z is the absolute magnitude and ϕ is the angle between $\tilde{Z}$ and R in Fig. 18-4. This angle is an absolute value in that it has no relationship to the axis of reference in a phasor diagram, as do voltage and current. Alternating current causes a voltage drop in resistance which is in time phase with the current and in inductive reactance a drop which leads the current by 90 electrical degrees, assuming the positive direction for measurement of angles is counterclockwise. Or conversely, the current in an inductive reactance lags the voltage drop by 90°.

Impedance Values. Tables are available which give 60-Hz impedance values in ohms per 1000 ft for common sizes of wire and cable. The values can be expressed in the form $\tilde{Z} = R + jX$, which can be converted to the form $Z\underline{/\phi°}$ if desired. The latter form is convenient to use in voltage-drop calculations when the current is expressed as $I\underline{/\phi°}$.

Power Factor. In typical distribution loads, the current lags the voltage, as shown in Fig. 18-3, where θ is shown as the angle between current and sending voltage and $\cos\theta$ is referred to as the *power factor* of the circuit. In a purely resistive circuit, the current and voltage are in phase; consequently, the power factor is 1.0 or unity. In a purely inductive circuit, the voltage and current are out of phase by 90 electrical degrees, resulting in a power factor of zero. In a circuit consisting of a resistance in series with a reactance of equal ohmic value ($\phi = 45°$), $\theta = \pm 45°$ also. Thus, the power factor is $\cos 45° = 0.707$, or 70.7%.

In a single-phase ac circuit, the load in kW can be expressed as

$$\text{kW} = EI \cos\theta \tag{18-5}$$

where E = magnitude of rms line-to-neutral voltage, kV
I = magnitude of current, rms amperes
θ = electrical angle between phasor voltage and current

From Eq. (18-5), it is obvious that the magnitude of the current for a given voltage and kilowatt load depends on the power factor, or

$$I = \text{kW}/(E \cos \theta) \tag{18-6}$$

The corresponding equations for balanced 3-phase circuits are

$$\text{kW} = \sqrt{3}\, EI \cos \theta \tag{18-7}$$

and

$$I = \text{kW}/(\sqrt{3}\, E \cos \theta) \tag{18-8}$$

where the symbols are as specified above, and θ is measured as the angle between the line-to-neutral voltage of a given phase and the current in that phase.

Example. Given a load of 500 kW at 80% power factor (lagging), 7.2 kV circuit voltage, 60-Hz, single-phase circuit using 1/0 aluminum conductor spaced 30 in on centers. The load is located 1 mi from the substation. What is the voltage drop? From tables on conductor characteristics,

$$r = 0.185\ \Omega/1000\ \text{ft}$$

$$x = 0.124\ \Omega/1000\ \text{ft}$$

Therefore, $$R + jX = 5.28\,(0.185 + j\,0.124) = 0.9769 + j\,0.6547\ \Omega$$

From Eq. (18-6),

$$I = \frac{\text{kW}}{E \cos \theta} = \frac{500}{7.2 \times 0.8} = 86.81\ \text{A}$$

$$E = 7.2\angle\theta^\circ$$

$$\cos \theta = 0.80$$

$$\theta = 36.87^\circ$$

and $$\sin \theta = 0.60$$

From Eq. (18-4),*

$$\text{Voltage drop} = 2(IR \cos \theta + IX \sin \theta) = (86.81 \times 0.9769 \times 0.8 + 86.81 \times 0.6547 \times 0.6)$$
$$= 2(67.84 + 34.10) = 203.88\ \text{V}$$

Calculation of 3-Phase Line Drops with Balanced Loads. In 3-phase circuits with balanced loads on each phase, the line-to-neutral voltage drop is merely the product of the phase current and the conductor impedance as determined from standard tables. There is no return current with balanced 3-phase loads. Thus, the line-to-line voltage drop is $\sqrt{3}$ times the line-to-neutral drop, or

$$V_{\text{drop}\,L-L} = \sqrt{3}(IR \cos \theta + IX \sin \theta) \tag{18-9}$$

*The factor of 2 is used for a single-phase system to represent the impedance of the outgoing conductor and the return conductor.

For example, assume that the circuit of the preceding example now is a 3-phase 12.47-kV circuit 1 mi long with the same 1/0 aluminum conductors at an equivalent spacing of 30 in and a load of $3 \times 500 = 1500$ kW at 0.8 pf lagging. What is the line-to-line voltage drop? R and X are the same values as previously; that is, $R + jX = 0.9769 + j\,0.6547\ \Omega$.

The current per phase from Eq. (18-7) is

$$I = \frac{\text{kW}}{\sqrt{3}\,E\cos\theta} = \frac{1500}{\sqrt{3}\times 12.47 \times 0.8} = 86.81\text{ A}$$

as before,

$$\begin{aligned} V_{\text{drop } L-L} &= \sqrt{3}\,(IR\cos\theta + IX\sin\theta) \\ &= \sqrt{3}\,(86.81 \times 0.9769 \times 0.8 + 86.81 \times 0.6547 \times 0.6) \\ &= 117.51 + 59.06 = 176.57\text{ V (approx.)} \end{aligned}$$

Calculation of Voltage Drop in Unbalanced Unsymmetrical Circuits. If there are n different wires $a, b, c, d, \cdots, n$ carrying currents $I_a, I_b, I_c, \cdots, I_n$, respectively, whether 2-, 3-phase, the voltage drop in wire a per mile at 60 Hz is

$$I_aR_a + j\left[0.2794\left(I_a\log\frac{1}{r} + I_b\log\frac{1}{D_{ab}} + I_c\log\frac{1}{D_{ac}} + \cdots + I_n\log\frac{1}{D_{an}}\right) + 0.03034\,\mu\,I_a\right] \quad \text{volts in phasor form} \qquad (18\text{-}10)$$

where currents are in phasor amperes, R_a is 60-Hz ohmic resistance of conductor a per mile, r is equivalent radius, in inches, of conductor a, D_{ab}, D_{ac}, and D_{an} are distances, in inches, between centers of conductors a and b, a and c, and a and n, and u is the permeability of conductor a (unity for nonmagnetic material). To get the drop in b, replace all a's by b's and all b's by a's in Eq. (18-10); similarly, to get the drop in c, interchange a's and c's; likewise for n. For 25 Hz, multiply that part of Eq. (18-10) which is in brackets by 25/60. Equation (18-10) gives voltage drop for any degree of load unbalance, power factor, or conductor arrangements. In using this formula, calculations are made easier by choosing voltage to neutral as the reference axis.

Approximate Method of Calculating Voltage Drop in Unbalanced, Unsymmetrical Circuits. Equation (18-10) requires laborious calculations and is used only when exact results are necessary. Voltage drops sufficiently accurate for engineering purposes can be calculated by using an equivalent impedance for each conductor. The reactance component of the equivalent impedance is computed from a spacing D equal to the geometric means of the interaxial distances of the other conductors to the conductor being considered. For instance, if there are four conductors a, b, c, and n for conductor a, $D = \sqrt[3]{D_{ab}, D_{ac}, D_{an}}$; for conductor b, $D = \sqrt[3]{D_{ab}, D_{bc}, D_{bn}}$.

Phasor and Connection Diagrams. Phasor and connection diagrams are drawn in computing voltage drops in unbalanced circuits. Figure 18-5 shows an unbalanced 4-wire 3-phase 4160Y/2400-V circuit with assumed loads, power factors, and equivalent line impedances. Phase-to-neutral drops between source and load are given by the following, using one of the many possible voltage-notation conventions:

$$\begin{aligned} V_{na} - V_{n'a'} &= I_aZ_a + I_nZ_n \\ V_{nb} - V_{n'b'} &= I_bZ_b + I_nZ_n \\ V_{nc} - V_{n'c'} &= I_cZ_c + I_nZ_n \end{aligned} \qquad (18\text{-}11)$$

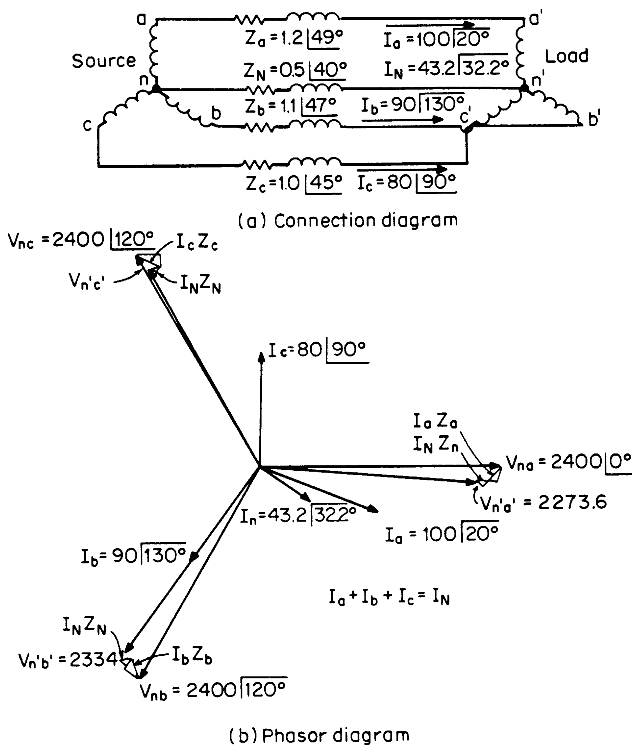


FIGURE 18-5 Connections and phasor diagrams for unbalanced loads and unsymmetrical circuit.

Phase-to-phase drops between source and load are given by the following:

$$
\begin{aligned}
V_{ba} - V_{b'a'} &= I_a Z_a - I_b Z_b \\
V_{ac} - V_{a'c'} &= I_c Z_c - I_a Z_a \\
V_{cb} - V_{c'b'} &= I_b Z_b - I_c Z_c
\end{aligned}
\tag{18-12}
$$

In computing line-to-neutral drop in phase a, it is convenient to choose V_{na} as the axis of reference.

$$
\begin{aligned}
V_{na} - V_{n'a'} &= I_a Z_a + I_n Z_n = (100\angle 20^\circ)(1.2\angle 49^\circ) + \mathrm{X}(43.2\angle 32.2^\circ)(0.5\angle 40^\circ) \\
&= 120\angle 29^\circ + 21.6\angle 7.8^\circ = 126.4 + j61.9
\end{aligned}
$$

Load voltage $V_{n'a'} = 2400 - 126.4 - j61.9 = 2273.6$ V (very nearly)

Likewise, in computing line-to-neutral drop in phase b, it is convenient to choose V_{nb} as the axis of reference. The phasor diagram of Fig. 18-5 must be rotated in a counterclockwise direction 120°; then $I_b = 90\angle 10^\circ$ and $I_n = 43.2\angle 87.8^\circ$.

$$
V_{nb} - V_{n'b'} = I_b Z_b + I_n Z_n = (90\angle 10^\circ(1.1\angle 47^\circ) + (43.2\angle 87.8^\circ)(0.5\angle 40^\circ) = 65.8 + j76.6
$$

Load voltage $V_{n'b'} = 2400 - 65.8 - j76.6 = 2334.2$ V (very nearly)

Drop in the neutral conductor of a 4-wire 3-phase circuit or a 3-wire 2-phase circuit makes resultant drop on the more heavily loaded phases greater than it would be for the same current under balanced conditions. Likewise, net drop is less on more lightly loaded phases than for the same current when balanced.

Distributed Loads, Voltage Drop, and I^2R Loss. Voltage drop and conductor power losses resulting from a concentrated load on a distribution line can be calculated easily as shown in earlier parts of this section. However, distribution circuit loads are generally considered to be distributed—often, but not always, uniformly. Distributed load may be considered as effectively concentrated at one point along the circuit to calculate total voltage drop and at another point to calculate conductor I^2R losses in the conductor. If the load is uniformly distributed along the feeder, the total voltage drop can be calculated by assuming that the entire load is concentrated at the midpoint of the circuit, and the total I^2R losses can be calculated by assuming that the load is concentrated at a point one-third the total distance from the source.

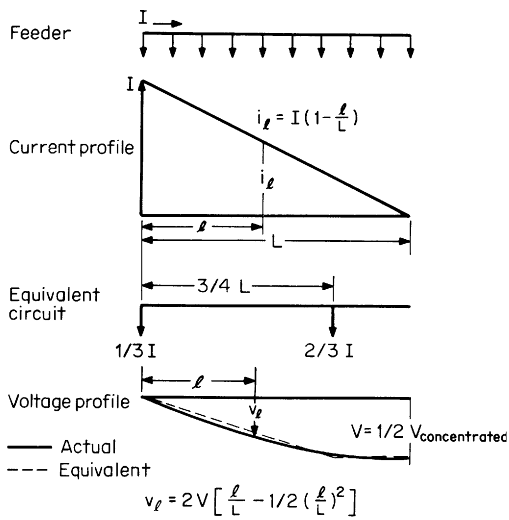


FIGURE 18-6 Uniformly distributed loads.

However, if there is a superimposed through load beyond the given feeder section, this method of calculation becomes cumbersome. It is possible to develop a single precise equivalent circuit for both the voltage-drop and loss calculations. Figure 18-6 shows the load representation and equivalent for uniformly distributed loads. Equivalents also can be developed for other types of distribution. Figure 18-6 shows the equivalent circuit of two-thirds of the total load concentrated at three-quarters of the total distance from the source.

18.5 THE SUBTRANSMISSION SYSTEM

Definition. *Subtransmission* is that part of the utility system which supplies distribution substations from bulk power sources, such as large transmission substations or generating stations. In turn, the distribution substations supply primary distribution systems. Subtransmission has many of the characteristics of both transmission and distribution in that it moves relatively large amounts of power from one point to another, like transmission, and at the same time it provides area coverage, like distribution.

In some utility systems, transmission and subtransmission voltages are identical; in other systems, subtransmission is a separate and distinct voltage level (or levels). This is easy to account for because in the evolutionary development of utility systems, today's transmission voltage naturally tends to become tomorrow's subtransmission voltage, just as today's subtransmission voltage tends to become tomorrow's primary distribution voltage.

Because of the wide range of voltages used in subtransmission, and because of the wide variation in geographic conditions and local ordinances, subtransmission circuits are sometimes built on pole lines on city streets, or on tower lines on private rights-of-way, or in underground cables.

Voltages. Voltages of subtransmission circuits range from 12 to 345 kV, but today the levels of 69, 115, and 138 kV are most common. The use of the higher voltages is expanding rapidly as higher primary voltages are receiving increased usage. Current practice as indicated by an informal utility survey is shown in Fig. 18-7; 115 and 138 kV together comprise about half the usage, 69 kV about 20%; 230 kV usage is becoming substantial, reflecting the growing use of 25- and 34.5-kV primary distribution.

Conductors of ACSR or aluminum generally have supplanted copper in overhead construction, and aluminum conductors are being used increasingly in cables.

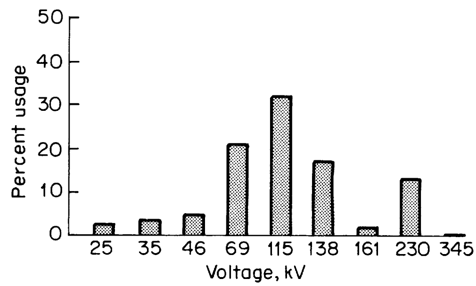


FIGURE 18-7 Use of distribution substation high-voltage rating.

Voltage Regulation of Subtransmission. The size of conductors used in subtransmission systems is determined by (1) magnitude and power factor of the load, (2) emergency loading requirements, (3) distance that the load must be carried, (4) operating voltage, (5) permissible voltage drop under normal and emergency loading, and (6) optimal economic balance between installed cost of the conductor and cost of losses. Table 18-2 gives the line-to-neutral voltage drops per 100,000 A · ft for common cable and overhead conductor sizes and representative power factors for 34.5- and 69-kV subtransmission. Values in the table are based on the approximate formula (18-4)

$$V_{\text{drop}} = IR \cos\theta + IX \sin\theta = IZ \cos(\phi - \theta)$$

where R, X, and Z are 60-Hz resistance, reactance, and impedance in ohms per 1000 ft of a single conductor, θ is the power-factor angle in electrical degrees, and ϕ is the impedance angle, $\tan^{-1}(X/R)$.

Examples of How to Use Table 18-2. Determine the voltage drop when a 3-phase 20,000-kVA load at 95% power factor is carried 10 mi over an overhead 69-kV circuit with No. 2/0 ACSR conductor. Assuming the receiving-end voltage to be 69 kV, the current is

$$I = \frac{\text{kVA}}{\sqrt{3}E} = \frac{20{,}000}{\sqrt{3} \times 69} = 167.35 \text{ A}$$

Circuit feet are

$$10 \times 5280 = 52{,}800 \text{ ft}$$

Thus

$$\frac{\text{A}\cdot\text{ft}}{100{,}000} = \frac{167.35 \times 52{,}800}{100{,}000} = 88.36$$

From the overhead portion of Table 18-2, the voltage drop per 100,000 A · ft at 95% power factor for a No. 2/0 ACSR conductor is 19.1 V. Therefore, the total voltage drop for the example is 88.36 × 19.1 = 1687.68 V line-to-neutral. Since normal line-to-neutral voltage is $69/\sqrt{3} = 39.838$ kV, or 39,838 V, the percent voltage drop is 1687.68 × 100/39,838 = 4.24%.

Assuming that permissible voltage drop is the limiting factor, what overhead ACSR conductor size should be used to supply a load of 40,000 kVA at 95% power factor and receiving-end voltage of 69 kV with a permissible drop of 5% and 8 mi between sending and receiving ends?

$$\text{Current} = \frac{40{,}000}{\sqrt{3} \times 69} = 334.71 \text{ A}$$

$$\text{Circuit feet} = 8 \times 5280 = 42{,}240 \text{ ft}$$

$$\frac{\text{A}\cdot\text{ft}}{100{,}000} = \frac{334.71 \times 42{,}240}{100{,}000} = 141.38$$

TABLE 18-2 Voltage Drops per 100,000 A · ft* for 3-Phase, 60-Hz, 34.5- and 69-kV Subtransmission

	Voltage class										
	34.5 kV					69 kV					Approx. amp. capacity for air moving at 2 ft/s
	Lagging power factor										
Conductor size	0.7	0.8	0.9	0.95	1.00	0.7	0.8	0.9	0.95	1.00	
Underground subtransmission†											
Aluminum:											
No. 1/0	18.3	19.9	21.1	21.5	21.0						
No. 2/0	15.4	16.5	17.4	17.6	16.9						
No. 4/0	10.7	11.2	11.5	11.4	10.5						
350 kcmil	7.69	7.84	7.77	7.55	6.50	8.04	8.10	7.92	7.62	6.38	
500 kcmil	6.15	6.12	5.88	5.59	4.50	6.53	6.43	6.10	5.74	4.48	
750 kcmil	4.96	4.80	4.44	4.10	3.00	5.25	5.05	4.63	4.23	3.01	
1000 kcmil	4.32	4.12	3.73	3.37	2.30	4.69	4.44	3.96	3.55	2.32	
Overhead subtransmission‡											
ACSR:											
No. 4	42.9	45.5	47.3	47.5	44.7	43.6	46.1	47.7	47.8	44.7	120
No. 2	31.5	32.5	32.7	32.1	28.4	32.2	33.1	33.1	32.4	28.4	165
No. 1/0	24.1	24.1	23.2	22.1	18.0	24.8	24.7	23.7	22.4	18.0	225
No. 2/0	21.6	21.2	20.1	18.8	14.6	22.3	21.8	20.5	19.1	14.6	260
No. 4/0	17.3	16.6	15.1	13.8	9.66	18.0	17.2	15.5	14.1	9.66	355
336.4 kcmil	12.7	11.8	10.4	9.13	5.57	13.4	12.4	10.8	9.44	5.57	480
477 kcmil	11.2	10.3	8.72	7.44	3.92	12.0	10.9	9.15	7.75	3.92	605
795 kcmil	9.73	8.68	7.06	5.78	2.37	10.4	9.28	7.49	6.09	2.37	850

Note: 1 in = 25.4 mm; 1 in^2 = 645 mm^2; 1 ft = 0.3048 m. Regulation of copper conductors can be estimated with reasonable accuracy as that of aluminum conductors two sizes larger. For ampacities of cables, see Tables 18–22 and 18–23.

*Values in the table give the difference in absolute value between sending-end and receiving-end line-to-neutral voltages of a balanced 3-phase circuit.

†Underground cable impedances are based on 90°C conductor temperature with close triangular spacing of cables using typical solid-dielectric insulation, 100% insulation level, single conductor, shielded and jacketed.

‡Overhead conductor impedances are based on 50°C conductor temperature, ACSR construction, 600 A/in^2 density with 60-in equivalent spacing for 35 kV and 90 in for 69 kV.

The permissible voltage drop is $0.05 \times 69{,}000/\sqrt{3} = 1991.92$ V line-to-neutral. The corresponding permissible voltage drop per 100,000 A · ft is

$$\frac{1991.92}{141.38} = 14.1 \text{ V}/100{,}000 \text{ A} \cdot \text{ft}$$

From Table 18-2 it is seen that this corresponds approximately to No. 4/0 ACSR.

Subtransmission System Patterns. A wide variety of subtransmission system designs are in use, varying from simple radial systems to systems similar to networks. The radial system is not generally used because most utilities today plan their subtransmission-distribution substation systems so that one major contingency such as outage of a subtransmission circuit or failure of a distribution substation transformer will not result in loss of load—or at least the loss of load will be of short duration while automatic switching operations take place. Thus, loop and multiple circuit patterns predominate. Figures 18-8 and 18-9 illustrate the basic nature of these two patterns. The loop pattern implies that a single circuit originating at one bulk power source "loops" through several substations before terminating at another bulk source or even at the original source. Reinforcing ties, as indicated by the dotted connection, are used when the number of substations exceeds some predetermined level.

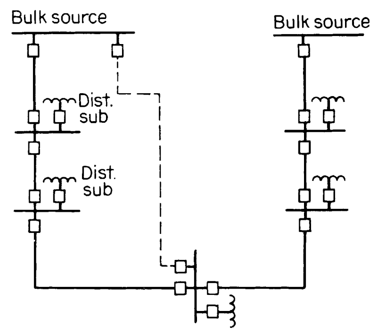


FIGURE 18-8 Loop pattern.

Multiple circuit pattern implies the use of two or more circuits which are tapped at each substation, as illustrated in Fig. 18-9. The circuits may be radial or may terminate in a second bulk power source. Many variations of the two basic patterns are found. From a recent informal survey of approximately 50 major utilities, it appears that the two patterns are about equally used.

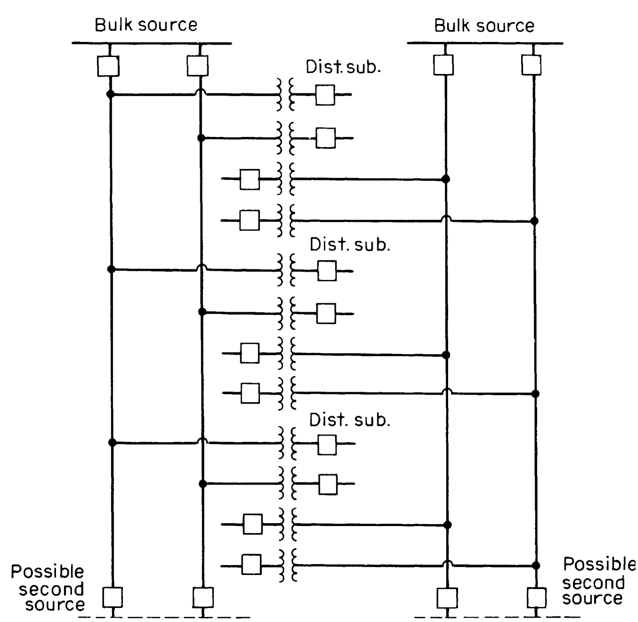


FIGURE 18-9 Multiple pattern.

A vast majority of today's subtransmission is of overhead construction, much of it built on city streets as contrasted with private rights of way. However, appearance and environmental considerations, difficulty in obtaining substation sites and rights of way, and rapid growth of underground distribution are certain to exert continuing pressure on the undergrounding of subtransmission. Even with the use of direct-buried, solid-dielectric cables, the cost of underground subtransmission is many times the cost of overhead circuits, particularly where the overhead subtransmission can be built on city streets.

Thus, a requirement to build future subtransmission underground would have major impact on the balance of overall subtransmission-substation-primary distribution costs. It undoubtedly would focus attention on minimizing the amount of subtransmission circuitry needed to cover the load area, which in turn would favor

Fewer, larger substations

Loop subtransmission pattern rather than multiple parallel circuits

Depending on load density in this area, it *could* favor

Higher primary voltage

Higher subtransmission voltage

Changes in either subtransmission or primary voltage levels are major decisions which require study in depth and ultimately the commitment of large financial resources.

18.6 PRIMARY DISTRIBUTION SYSTEMS

The primary distribution system takes energy from the low-voltage bus of distribution substations and delivers it to the primary windings of distribution transformers.

Overhead Primary Systems. Typically, overhead primary distribution systems have been operated as radial circuits (normally open loops) from the substation outward. Figure 18-2 shows schematically a typical primary feeder in a predominantly residential area; an overhead 12.47Y/7.2-kV system is used for illustrative and functional purposes, but underground systems will be discussed later.

The main feeder backbone usually is a 3-phase 4-wire circuit from which the single-phase lateral or branch circuits are tapped through fuse cutouts to protect the system from faults on the lateral circuits. The single-phase lateral circuits consist of one phase conductor and the neutral. Distribution transformers are connected between the phase and the neutral; in this case they would have a rating of 7200 V.

Utilities use automatic reclosing feeder breakers and line reclosers to minimize service interruptions. However, serious problems involving the main will cause an outage to some or all of the feeder until line crews can locate the problem and manually operate pole-top disconnecting switches appropriately to isolate the problem and to pick up as much load as possible from adjacent feeders. Switches of this kind usually are found in both the main and lateral circuits, as indicated in Fig. 18-2. Also, it is often possible to make and to remove connections while the system is energized through the use of hot-line tools, hot-line clamps, insulated bucket trucks, etc.

Generally, this approach has provided an acceptable level of service because overhead system troubles are relatively easy to locate, and repair times are short. However, when the entire primary system is installed underground, while the frequency of serious trouble is expected to be lower than in overhead systems, it is likely that the time involved in pinpointing the location and making repairs will be much longer than in overhead systems.

Underground System. While a relatively small percentage of new general-purpose feeders is being installed totally underground, the trend is growing and is expected to continue to grow.

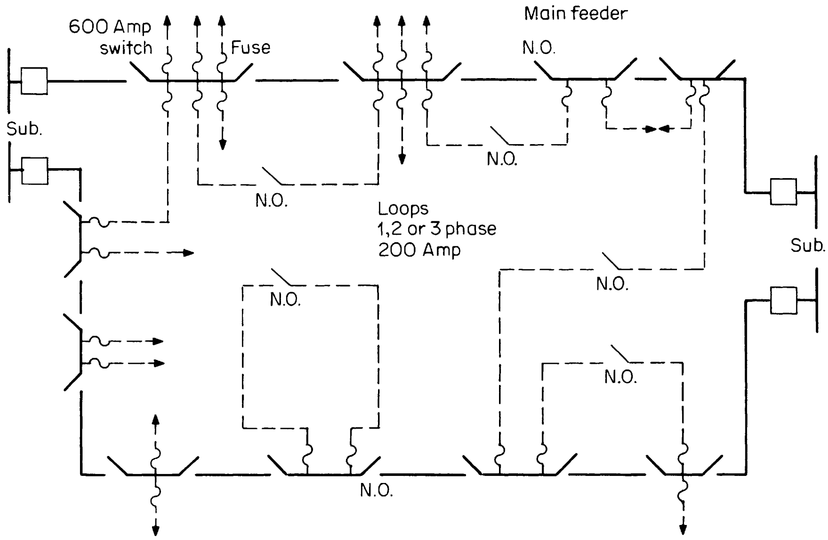


FIGURE 18-10 Typical main-feeder underground circuit. (All switches closed unless shown otherwise.)

Since it is difficult to accomplish many maintenance and operating functions on an underground system while it is "hot," or energized, in contrast to overhead-system practices, specific provisions must be made in the system design to incorporate needed sectionalizing and overcurrent protective equipment.

The main feeder plan shown in Fig. 18-10 is reasonably typical of present practice on underground systems supplying basically residential and small commercial loads. Note that the main feeders are operated radially, but with normally open ties to adjacent main feeders. The main feeder switches usually are 3-phase, 600-A, manually operated load-break switches. The single-phase and 3-phase lateral circuits also are operated as normally open loops.

Switching in the 200-A circuits can be accomplished by means of either load-break switches or separable, insulated cable connectors. Usually, two main feeder switches are grouped along with the lateral circuit switching and protective equipment into one piece of pad-mounted equipment.

The primary feeders supplying secondary-network systems in metropolitan areas usually are radial 3-wire circuits consisting of 3/c cables in underground duct lines. The 3-phase network transformers are T-tapped to the primary feeders.

Automation. With increasing emphasis on reliability of service, a definite trend is under way to make greater use of protective and sectionalizing equipment in the primary system in order to minimize the number of customers involved in an outage and to reduce the outage time. Proposed schemes run the gamut from manually operated devices to automatic devices remotely controlled from distribution centers. The remote-controlled schemes vary from some type of supervisory control to computer-controlled systems with built-in logic to cope quickly with the various problems which may arise.

Primary-Distribution-System Voltage Levels. Since World War II, the 15-kV distribution class has become firmly entrenched and today represents 60% to 80% of all primary distribution activity. Very little expansion of lower-voltage systems is taking place. There is a trend, however, toward increasing usage of primary voltage levels above the 15-kV class. This trend has an impact on substation and subtransmission practices as well because higher primary voltages almost axiomatically lead to larger substations and higher subtransmission voltages.

The two principal voltages above 15 kV are 24.49Y/14.4 kV and 34.5Y/19.92 kV. New line additions at these voltage levels now average more than 20% of those at 15 kV.

To achieve economy, the higher primary voltages also require heavier feeder loadings which could imply reduced service reliability because more customers are affected by primary faults. Greater use of automatic switching and protective equipment can do much toward preserving a level of reliability to which the public has become accustomed. This is another reason that most observers believe that an increased amount of automation is inevitable in our distribution systems.

For example, a typical 12.47-kV feeder serves a normal peak load on the order of 6000 to 7000 kVA. On this basis, the probable peak loading of a fully developed 34.5-kV feeder would be expected to be in the neighborhood of 18,000 to 20,000 kVA.

Why go to high-voltage distribution (HVD)? Most of today's systems in the 15-kV class are not voltage-drop-limited, and cost of higher-voltage laterals and associated equipment needed to cover the load area is greater. The major economic advantages are:

1. Larger (and fewer) substations
2. Fewer circuits
3. Possibility of eliminating a system voltage-transformation level where the new primary voltage is the former subtransmission level

Other advantages of HVD which are difficult to evaluate in dollars are:

1. Reduced losses in early stages of development
2. Reduced voltage regulation
3. Greater distance or area coverage
4. Fewer circuits per route (reduced congestion)
5. Fewer circuit positions at substations
6. Fewer substation sites
7. Greater flexibility in supplying large spot loads

Some of the disadvantages of HVD have been

1. Cost of equipment
2. Reliability due to increased exposure
3. Higher equipment failure rates
4. Operability

Conductor Sizes. The conductor sizes used in overhead primaries generally range from No. 2 AWG to 795 kcmil. ACSR and aluminum conductors have almost entirely displaced copper for new construction. Aerial cable is used occasionally for primary conductors in special situations where clearances are too close for open-wire construction or where adequate tree trimming is not practical. The type of construction more frequently used consists of covered conductors (nonshielded) supported from the messenger by insulating spacers of plastic or ceramic material. The conductor insulation, usually a solid dielectric such as polyethylene, has a thickness of about 150 mils for a 15-kV class circuit and is capable of supporting momentary contacts with tree branches, birds, and animals without puncturing. This type of construction is commonly referred to as *spacer cable*.

The conductor sizes most commonly used in underground primary distribution vary from No. 4 AWG to 1000 kcmil. Four-wire main feeders may employ 3- or 4-conductor cables, but single-conductor concentric-neutral cables are more popular for this purpose. The latter usually employ crosslinked polyethylene insulation, and often have a concentric neutral of one-half or one-third of the main conductor cross-sectional area.

The smaller-sized cables used in lateral circuits of URD systems are nearly always single-conductor, concentric-neutral, crosslinked polyethylene-insulated, and usually directly buried in the earth. Insulation thickness is on the order of 175 mils for 15-kV-class cables and 345 mils for 35-kV class with 100% insulation level.

Stranded or solid aluminum conductors have virtually supplanted copper for new construction, except where existing duct sizes are restrictive. With the solid-dielectric construction, in order to limit voltage gradient at the surface of the conductor within acceptable limits, a minimum conductor size of No. 2 AWG is common for 15-kV-class cables, and No. 1/0 AWG for 35-kV class.

Voltage Regulation of Primary Distribution. Table 18-3 can be used to determine the voltage drop of an existing circuit when the load data are known or to determine minimum conductor size required to meet a given voltage-drop limit. Data are given for various underground-cable and overhead-conductor configurations for 12.47 and 34.5 kV.

Example. What is the voltage drop for a 34.5-kV overhead circuit 3 mi long using 4/0 aluminum conductor and carrying a balanced 3-phase load of 15,000 kVA at 90% power factor: The current is $15{,}000/\sqrt{3} \times 34.5 = 251$ A. The circuit feet are $3 \times 5280 = 15{,}840$ ft. Thus $\text{A} \cdot \text{ft}/100{,}000 = 251 \times 15{,}840/100{,}000 = 39.758$. From Table 18-3, the appropriate voltage drop per 100,000 A · ft is 14.0 V line-to-neutral. Therefore, the total voltage drop for the example is

$$39.758 \times 14.0 = 556.6 \text{ V line-to-neutral}$$

Since normal line-to-neutral voltage is $34{,}500\sqrt{3} = 19{,}920$ V, the percent voltage drop is

$$556.6 \times 100/19{,}920 = 2.79\%$$

Example. What is the minimum aluminum conductor size to carry 6000 kVA at 90% power factor of balanced 3-phase load over a 2-mi, 12.47Y/7.2-kV feeder with no more than a 3% voltage drop? Load current is $6000/\sqrt{3} \times 12.47 = 277.8$ A. Circuit feet $= 2 \times 5280 = 10{,}560$ ft. Thus

$$\frac{\text{A}\cdot\text{ft}}{100{,}000} = \frac{277.8 \times 10{,}560}{100{,}000} = 29.34$$

$$\text{Permissible voltage drop} = 0.03 \times \frac{12{,}470}{\sqrt{3}} = 216 \text{ V}$$

The corresponding drop per 100,000 A · ft is $216/29.34 = 7.36$ V, line-to-neutral. From Table 18-3, this value falls between 477 and 795 kcmil, so that the latter size would be chosen.

Loading. Loading of primary feeders varies greatly depending on primary voltage, load density, emergency loading requirements, etc. Typical peak loads on 15-kV class feeders are 6 to 7000 kVA. Peak loads on 25- and 35-kV class, fully developed feeders probably will be proportionally greater in the future, assuming that appropriate measures can be taken to maintain acceptable reliability of service.

Voltage Drop. Voltage drop in the primary feeder is an important factor in system design; however, it is only one of the many voltage-drop considerations involved in determining the range of voltages delivered to the customers' service entrances. American National Standard, "Voltage Ratings for Electric Power Systems and Equipment (60-Hz)," ANSI C84.1-1995 (R200), defines in detail the voltage ranges which should be observed. Outside the distribution substation, voltage drops occur in the primary system, the distribution transformer, the secondary system, the service drop, and in the users' wiring systems as well. Remedial measures, such as voltage regulators and shunt capacitor banks, can be used to counteract or reduce the voltage drop due to load flow.

A traditional rough rule of thumb has been to allow a voltage drop of about 3% in the primary of urban and suburban systems at time of peak load. Actually, with typical load densities and primary systems of 15-kV class or higher, it is very probable that economic system designs have a primary voltage drop smaller than 3%.

TABLE 18-3 Line-to-Neutral Voltage Drops per 100,000 A · ft* for 12.47Y/7.2 and 34.5Y/19.92 kV and Balanced 3-Phase Loads

	Voltage class										
	12.47Y/7.2 kV					34.5Y/19.92 kV					Approx. amp. capacity for air moving at 2 ft/s
	Lagging power factor										
Conductor size	0.7	0.8	0.9	0.95	1.00	0.7	0.8	0.9	0.95	1.00	
Underground primary											
Aluminum:											
Concentric neutral—direct buried, cross-linked polyethylene, conductor 70°C, neutral 60°C, earth resistivity 90 Ω · cm^3, triplex configuration, full installation											
No. 1/0	17.1	18.5	19.8	20.2	19.8	17.6	19.0	20.1	20.4	19.8	
No. 2/0	14.1	15.1	16.0	16.3	15.7	14.6	15.6	16.3	16.5	15.7	
No. 4/0	9.82	10.4	10.7	10.7	9.96	10.3	10.8	11.0	10.9	9.95	
350 kcmil	7.01	7.19	7.17	7.00	6.11	7.37	7.49	7.39	7.16	6.11	
500 kcmil	5.66	5.69	5.55	5.31	4.40	6.04	6.00	5.76	5.47	4.40	
750 kcmil	4.63	4.55	4.30	4.03	3.12	4.95	4.82	4.49	4.16	3.11	
1000 kcmil	4.10	3.98	3.69	3.41	2.52	4.37	4.20	3.85	3.52	2.51	
Single conductor shielded and jacked, cross-lined polyethylene, conductor 70°C, unigrounded shield, triplex configuration, full insulation											
350 kcmil	7.29	7.49	7.51	7.35	6.47	7.55	7.72	7.67	7.47	6.47	
500 kcmil	5.78	5.82	5.67	5.45	4.54	6.08	6.07	5.86	5.58	4.54	
750 kcmil	4.64	4.54	4.26	3.97	3.02	4.88	4.74	4.41	4.08	3.02	
1000 kcmil	4.02	3.85	3.52	3.21	2.26	4.23	4.03	3.65	3.31	2.26	
Overhead primary†											
No. 4	42.3	45.4	47.8	48.5	46.6	43.4	46.3	48.5	49.0	46.6	115
No. 2	29.8	31.2	32.0	31.9	29.3	30.9	32.2	32.7	32.4	29.3	160
No. 1/0	21.8	22.2	22.1	21.5	18.5	23.0	23.2	22.8	22.0	18.5	215
No. 2/0	19.0	19.1	18.6	17.8	14.7	20.1	20.0	19.3	18.3	14.7	250
No. 4/0	14.7	14.3	13.3	12.4	9.20	15.9	15.3	14.0	12.7	9.20	340
336.4 kcmil	11.8	11.2	9.97	8.91	5.80	13.0	12.1	10.7	9.41	5.80	465
477 kcmil	10.4	9.58	8.27	7.18	4.10	11.5	10.5	8.97	7.68	4.10	590
795 kcmil	8.22	7.92	6.52	5.40	2.40	9.96	8.88	7.22	5.90	2.40	820

Note: 1 in = 25.4 mm; 1 ft = 0.3048 m. For ampacities of cables, see Tables 18-23 and 18-24. Regulation of copper for overhead conductors can be estimated with reasonable accuracy the same as that of aluminum conductors two sizes larger. For single-phase overhead primaries, the voltage drop is approximately two times the 3-phase values given in the table. For underground single-phase primaries in concentric-neutral, direct-buried cables, see section on URD systems. Cables are 15- and 35-kV classes, respectively.

*Values in the table give the difference in absolute value between sending-end and receiving-end line-to-neutral voltages of a balanced 3-phase circuit, in volts.

†Overhead conductor impedances are based on 50°C conductor temperature, aluminum conductor with 30-in equivalent spacing for 12.47Y kV and 60-in for 34.5Y kV.

In rural systems which are typified by long lines and light load densities, primary voltage drops may be somewhat larger. This is offset somewhat by the absence of secondaries in serving individual farms; however, the service drops often are longer than in urban systems. The design objective, of course, is to keep delivered voltage to all customers in an acceptable and satisfactory range.

18.7 THE COMMON-NEUTRAL SYSTEM

The 4-wire, multigrounded, common-neutral distribution system now is used almost exclusively because of the economic and operating advantages it offers. Usually, the windings of the substation transformers serving the primary system are wye-connected, and the neutral point is solidly grounded. Occasionally, a small amount of impedance is connected between the transformer neutral and ground in order to limit line-to-ground short-circuit currents on the primary system to a predetermined value. The neutral circuit must be a continuous metallic path along the primary routes of the feeder and to every user location. Where primary and secondary systems are both present, the same conductor is used as the "common" neutral for both systems. The neutral is grounded at each distribution transformer, at frequent intervals where no transformers are connected, and to metallic water pipes or driven grounds at each user's service entrance. The neutral carries a portion of the unbalanced or residual load currents for both the primary and secondary systems. The remainder of this current flows in the earth and/or the water system. For typical conditions, it is estimated that about one-half the return current flows in the neutral conductor, although the division can vary widely depending on earth resistivity and the relative routing of the electric and water systems. Figure 18-11 is a schematic representation of a common-neutral system.

Grounding of Neutral. Rules related to grounding on the utility system neutral are given in the National Electrical Safety Code (NESC), ANSI C2, and regulations governing the grounding of the neutral on users' premises are stated in the National Electrical Code (NEC), NFPA 70. In brief, the secondary neutral is grounded at every service through a metallic water-piping system and through "made electrode grounds" such as other underground metal systems, building steel, or driven ground electrodes. The increasing use of nonmetallic water piping and insulating couplings on metal water systems is requiring the use of other grounding means. The secondary neutral also is grounded at the distribution transformer, usually by means of driven grounds. Although it is often

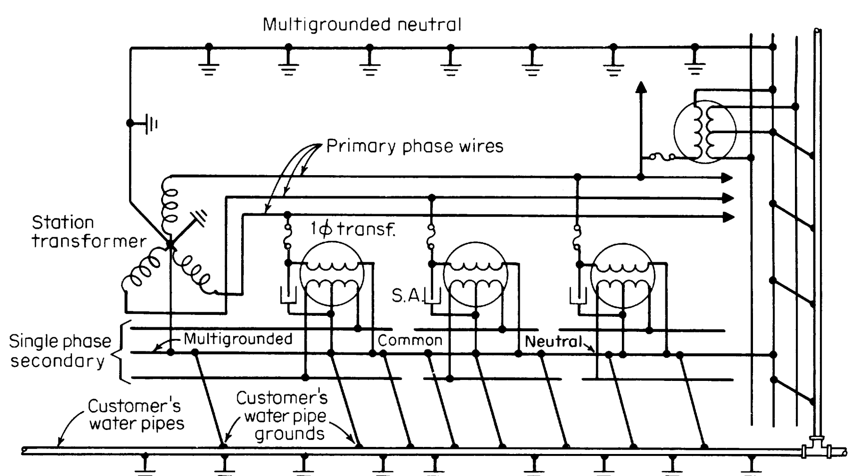


FIGURE 18-11 Common-neutral methods of distribution.

general practice to install a metal butt plate or a wire butt wrap on poles to help in grounding the system neutral and other equipment, the NESC requires two such devices to equal one *made electrode*; as a result, neither can be used to satisfy the NESC requirement for a direct earth ground with a made electrode at each transformer or other arrester location.

The resistance to ground of a typical metallic water-piping system usually is less than 3 Ω. When made electrode grounds are used, they should have a resistance of not more than 25 Ω. Many utilities strive for lower values such as 5, 10, or 15 Ω.

Where there is no secondary neutral as such and no distribution transformers, the primary neutral should be grounded at intervals of not less than 1000 ft. Many utilities require grounding at smaller spacing, such as 500 ft; to meet the NESC requirements for a multigrounded neutral, there must be a minimum of the equivalent of four made electrodes in *each* mile. In URD systems, the primary circuits usually are in direct-buried, concentric neutral cable, so that excellent grounding is obtained.

The neutral must have a continuous metallic path between the substation and users' services. No disconnecting devices should be installed in the common neutral. In no case should the earth or buried metallic-piping systems be used as the only path for the return of normal load current.

Size of Primary Neutral. On single-phase primary circuits (phase and neutral), the neutral conductor should be large enough to carry almost as much current as the phase conductor. Often the same neutral conductor size is used for both, or the neutral has "100%" conductivity.

In 3-phase primary circuits carrying reasonably balanced load, the neutral conductor can be considerably smaller than the phase conductors; 50% conductivity is not uncommon; some utilities specify size of neutral conductor, such as No. 1/0 aluminum, regardless of the size of the phase wires.

Secondary-system neutral conductors are often the same size as the phase conductors where open-wire construction is used. Where triplexed construction is used, the neutral frequently has a reduced cross section.

4-Wire vs. 3-Wire Systems. The 4-wire, common-neutral primary system has many advantages over 3-wire systems:

1. Single-phase branch circuits, or laterals, consist of one insulated phase conductor and the neutral, rather than two insulated phase conductors. The economic advantage is very great in underground systems.
2. On overhead systems, only one lightning arrester is required at each single-phase distribution transformer, rather than two.
3. Only one primary bushing or cable termination is needed on each single-phase distribution transformer, rather than two. In the case of underground systems where the primary "loops through" each distribution transformer, two primary cable terminations or connectors are needed, rather than four.
4. Only one fuse or fuse cutout is needed in the primary of each single-phase distribution transformer. Not only is this a substantial economic advantage, but a short circuit in the primary of the transformer is interrupted positively by the action of a single fuse, and primary voltage is thereby removed from the transformer. In the case of the 3-wire system with the distribution transformer connected phase-to-phase, a second fuse must operate to remove primary voltage and the fault. There may be appreciable time between operation of the two fuses during which fault current continues to flow and abnormal voltages may be experienced by the user.
5. Single-phase primary lateral circuits can be protected by a single fuse cutout, rather than two. Line-to-ground short circuits are promptly cleared by operation of one fuse and voltage removed from the branch circuit. In a 3-wire system (assumed grounded at the substation), single-phase lateral protection, if used, would require two fuse cutouts; a line-to-ground fault would blow only one fuse, leaving all the distribution transformers on that circuit excited at only 58% of normal as long as the faulted phase remains grounded. Under these conditions users' equipment would be exposed to abnormally low voltage. The ability to fuse lateral circuits contributes substantially to

reliability of service, since a major amount of the total circuit exposure comprises the primary laterals in residential areas.

Common-Neutral and Telephone Circuits. Usually, no problems are encountered in the joint use of poles for overhead distribution circuits and telephone circuits, particularly when the telephone circuits are in cable, as is now common practice. Also, in underground residential circuits, power cables and telephone cables often are installed in the same trench with no intentional physical separation of the power and communication facilities, that is, "random lay." Where separate grounding electrodes are employed for supply and communication facilities at customer's premises, the electrodes shall be bonded together with not less than No. 6 AWG copper wire.

18.8 VOLTAGE CONTROL

System Voltage Levels and Voltage Ranges. Since about 1900, there have been several recommendations for certain voltages as standard or preferred for primary and secondary distribution systems, as well as for higher-voltage systems. The latest listing of standard system voltages is American National Standards Institute (ANSI) Standard C84.1-1995(R200), "Voltage Ratings for Electric Power Systems and Equipment (60 Hz)." This standard was formulated by both utilities and manufacturers, and its recommendations are followed by both segments of the industry. Observance of this standard enables the utilities and manufacturers to work in harmony. In many states, ANSI C84 is the basis for rulings of the regulatory commission as far as voltage requirements are concerned.

This standard designates certain standard nominal voltages, including 120/240 V single-phase, 480Y/277 V, 12,470Y/7200 V, as well as the higher primary voltages, 24,940Y/14,400 V and 34,500Y/19,920 V, and others.

Using the nominal 120/240-V system as an example, the standard designates two different ranges of voltage, range A and range B. Range A service voltage specifies that a utility supply system be so designed and operated that most service voltages are within the limits specified, for example, 114/228 and 126/252 V. The occurrence of service voltages outside these limits is to be infrequent.

With the typical voltage drops between the service entrance and the points of utilization, the utilization equipment is designed and rated to give fully satisfactory performance within range A.

Range B service voltage includes voltages above and below range A that necessarily result from practical design and operating conditions on supply or user systems. These conditions are limited in extent, frequency, and duration. When they occur, corrective measures should be undertaken within a reasonable time to improve voltages to meet range A requirements.

Insofar as practicable, utilization equipment is designed to give acceptable performance within range B. The design and operating bogey of the utilities is to provide service voltage to all customers at all times within range A limits.

Voltage Profiles. It is usually convenient to discuss distribution-feeder-voltage regulation in terms of voltage *profiles* of the feeder, because the voltages are everywhere different on the feeder. A profile is simply a graph of feeder-voltage magnitude versus location on the feeder. For a simple case of one load at the end of the feeder (assuming uniform conductor), the one-line diagram and profile are as shown in Fig. 18-12.

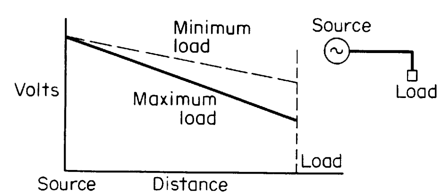


FIGURE 18-12 Voltage profile for concentrated load.

The profile is a straight line between source and the load, and the voltage regulation at any point between is proportional to the distance from the source. It may be, as shown by the dashed-line profile, that minimum load is not zero, in which case the voltage variation is less than the calculated

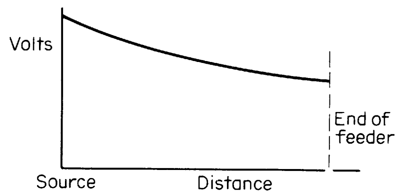


FIGURE 18-13 Voltage profile for distributed load.

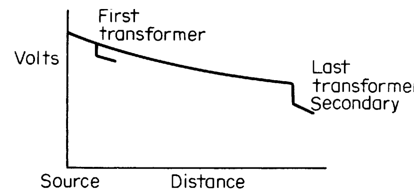


FIGURE 18-14 Additional regulation due to transformer and secondary.

regulation, since regulation is usually calculated on the voltage difference between no-load and full-load conditions. If additional loads are distributed along the feeder, the profile becomes a broken line, and if the load is uniformly distributed, the profile becomes a smooth curve, as shown in Fig. 18-13.

The shape of the profile is of less consequence than knowing the extremes, because there are generally customers connected at all points on the feeder, and no customer's voltage should be too high or too low. Since most feeders neither supply a single load nor are uniformly loaded, it usually is necessary to calculate the voltage profile on a piece-by-piece basis, representing the loads and feeder configurations as accurately as the situation warrants.

In addition to the distribution-feeder-voltage profile, there is additional regulation in the distribution transformer and its secondaries and services. This additional regulation can be added to the profile as shown in Fig. 18-14. For protection of the first customer on the feeder 0 from possible overvoltage, it is usual to assume only a partially loaded transformer rather than one at full load.

It is now possible to establish a limiting band of voltage within which all customers must lie for satisfactory service, usually range A. In turn, this also will establish the maximum permissible difference between the full-load and light-load primary voltage. The problem of holding the right voltage at each customer location at all times may be visualized by referring to Fig. 18-15.

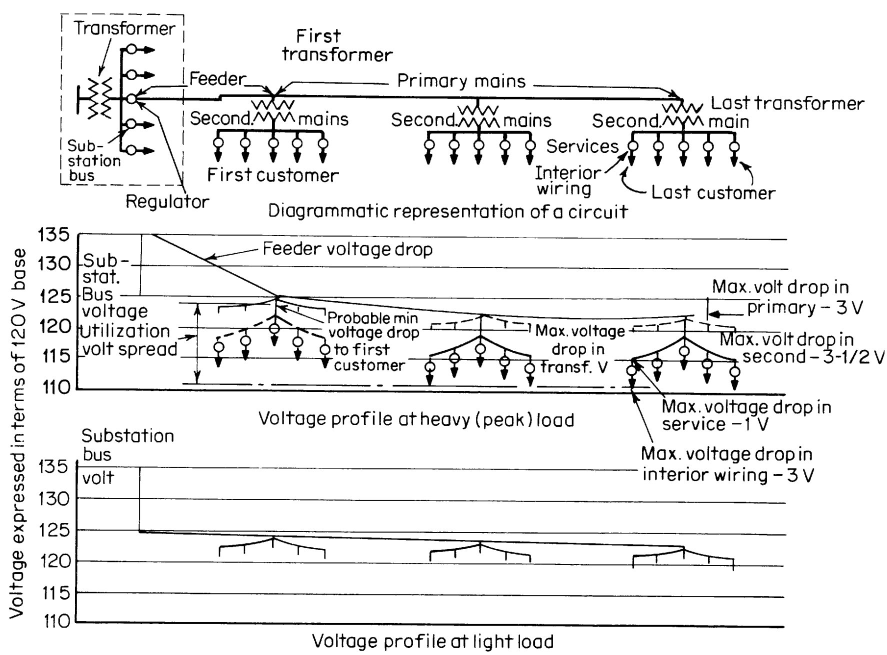


FIGURE 18-15 Distribution circuit with voltage profiles at heavy and light loads.

Voltage Control. As implied in Fig. 18-15, usually there is voltage control equipment in the substation consisting of load-tap changers on the power transformers or bus or feeder voltage regulators. This regulating equipment can control only the voltage *level* of the primary system. It can have no effect on the voltage *spread* between the first and last customers on the feeder.

There are several procedures which can be taken to correct for increasing voltage drops as the load on the feeders grows; among them are capacitors and supplementary feeder-voltage-regulator installations.

The effect of capacitor application is illustrated in Fig. 18-16, where the load is assumed to be uniformly distributed along the feeder, and a capacitor bank is installed as indicated. The capacitor produces a voltage rise because of its leading current flowing through the inductive reactance of the feeder. As is seen in the figure, this voltage rise increases linearly from zero at the substation to its maximum value at the capacitor location. Between the capacitor location and the remote end of the feeder, the rise due to the capacitor is at its maximum value.

When the capacitor voltage-rise profile is combined with the original feeder profile, the resulting net profile is obtained. The capacitor has increased the voltage *level* all along the feeder, resulting also in a reduced voltage *spread*.

In practical applications, the capacitor bank can be a permanently connected or "fixed" bank as shown or an automatically switched bank. The fixed bank is limited in size by the allowable voltage rise during light-load conditions, and therefore may not produce sufficient voltage rise during heavy-load conditions. It can be supplemented by additional switched capacitors which automatically switch on at heavy-load conditions and off again as the load decreases.

The effect of applying a supplementary feeder-voltage regulator is shown in Fig. 18-17. Note that the regulator produces no voltage effect between the source and the regulator location and its entire boost effect is between the regulator location and the remote end of the feeder.

A typical primary feeder serves distributed loads, as well as concentrated loads, and may also have shunt capacitors and supplementary voltage regulation, such that all these previous concepts must be employed in studying voltage conditions.

Voltage Regulation. Voltage regulation in distribution substations usually is accomplished by individual feeder-voltage regulators or by automatic load-tap-changing equipment in the substation transformers. Individual feeder-voltage regulators are advantageous where feeders of differing lengths and diverse load characteristics are supplied from the same substation bus. Automatic load-tap-changing equipment in the power transformer provides voltage control on the substation bus, or group regulation, when feeder lengths and load characteristics are reasonably homogeneous.

Voltage control is needed to compensate not only for the voltage regulation in the subtransmission system and substation transformer, which is measurable at the substation, but also for the voltage regulation which occurs in the distribution transformers and in the primary and secondary systems beyond the substation. The latter portion of the overall system voltage regulation is a function

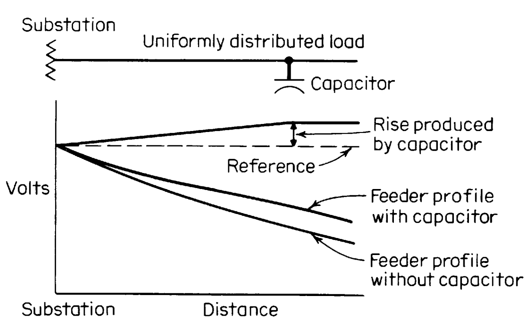


FIGURE 18-16 Effect of shunt-capacitor application.

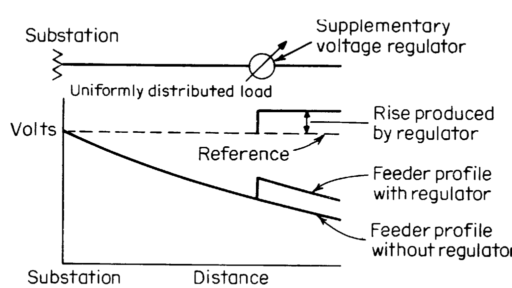


FIGURE 18-17 Effect of supplementary voltage regulator.

of the load flow and system impedances and cannot be measured directly at the substation. Therefore, the control systems of the voltage regulators or tap-changing equipment not only sense the voltage at the substation but also usually contain a "line-drop compensator" which simulates the voltage drop between the station and some point in the distribution system and controls the regulating equipment accordingly. Switched shunt capacitor banks sometimes are installed at the distribution substation as part of the overall system voltage control.

Feeder-Voltage Regulator. In the typical radial primary system, it is often necessary to regulate the voltage of each feeder separately by means of feeder-voltage regulators. These regulators may be of single-phase or 3-phase construction. The former are available in sizes from 25 to more than 400 kVA, the latter from 500 to 2000 kVA. For distribution-system application they are commonly available for voltages from 2.5 kV to 34.5 kV grd Y. Regulators commonly are capable of raising or lowering the voltage delivered to the feeder by 10% and normally are rated on this basis.

Modern voltage regulators all are of the step-voltage type, which has completely supplanted the earlier induction-voltage regulators. The step-voltage regulator basically is an autotransformer which has numerous taps in the series winding. Taps are charged automatically under load by a switching mechanism which responds to a voltage-sensing control in order to maintain voltage as close as practicable to a predetermined level. The voltage-sensing control receives its inputs from potential and current transformers and provides control of system voltage level and bandwidth. In addition, it permits selection of line-drop compensation and provides features such as operation counter, time-delay selection, test terminals, and control switch.

Most feeder-voltage regulators are of the 32-step design. Since they usually operate over a range of voltage of 20%, the voltage change per step is $^5\!/_8$%. If the full range of regulation of $\pm 10\%$ is not required, the regulators can carry more than rated current. For example, operating with a range of $\pm 5\%$, 160% of rated current can be carried.

Line-Drop Compensator. In simplified terms, the regulator voltage (local voltage) is stepped down by means of a potential transformer and fed to the control system, where it is compared with the desired and preset voltage level. If the actual voltage deviates from the preset level by more than $\pm\frac{1}{2}$ of the bandwidth, which also is preset by the operator, the tap-changing mechanism operates, after a preset time delay, to return the voltage within the preset band. From a practical point of view, the minimum bandwidth is twice the size of the voltage step, or $2 \times {}^5\!/_8\% = 1.25\%$. Maintaining a small bandwidth is important in reducing voltage variations and in making full use of the allowable system voltage drop.

The line-drop compensator consists of adjustable resistance and reactance components and is preset to simulate system impedance. By means of a current transformer, current proportional to load current is circulated through the resistance and reactance, producing a voltage signal which is combined with the signal from the local voltage. The net result is that the line-drop compensator causes a higher voltage to be held at the voltage regulator during periods of heavy load. In this way, a constant voltage is held at some point in the system, as determined by the compensator setting. This helps to achieve the goal of minimizing the voltage change with varying loads at any location.

Supplementary Voltage Regulation. In some long primary circuits, such as rural feeders, it is often necessary to provide voltage regulation in addition to that incorporated in substation equipment because of large voltage drops in the system. This supplementary voltage regulation usually is improved by single-phase automatic step regulators in the smaller ratings. These regulators are suitable for pole mounting.

Bus Regulation. Bus regulation at the distribution substation usually is provided by automatic load-tap-changing equipment built into the substation transformer or by large step-voltage regulators.

Switched Shunt Capacitors. Switched shunt capacitors are often applied at distribution substations or out on the primary feeders to accomplish a portion of the overall voltage-regulation job. Most utilities apply shunt capacitors primarily as a tool in economic system design. Usually fixed (unswitched) shunt capacitors are applied to bring the light-load power factor to more or less 100%. Then, additional automatically switched shunt capacitor banks are added to achieve an economic full-load power factor, which is usually in the order of 95% to 100%.

These capacitors, in addition to their economic functions, such as reducing losses and releasing system capacity, improve system conditions substantially. Usually additional voltage control is needed, however, and this is most economically accomplished with voltage-regulating equipment.

18.9 OVERCURRENT PROTECTION

General Principles. Coordination of overcurrent protection devices means their proper arrangement in series along a distribution circuit so that they function to clear faults from the lines and equipment in accordance with a prearranged sequence of operation. Fuse cutouts, automatic circuit reclosers, sectionalizers, and relayed circuit breakers are the overcurrent protective devices most commonly used. Ratings and characteristics can be obtained from appropriate product bulletins of the manufacturers.

When the protective devices are properly applied and coordinated:

They can eliminate service outages resulting from temporary faults.

They reduce the extent of outages, that is, the number of users affected.

They are helpful in locating the fault, thereby reducing the duration of interruptions.

Main-Line Sectionalizing. Usually, the first protective device on a primary feeder is a circuit breaker or a power-class recloser located in the substation. If the circuit is overhead, the circuit breaker often is provided with reclosing relays so that it operates in much the same manner as a recloser. If the circuit is primarily underground, reclosing is not generally used.

If portions of the main feeder and long branches extend beyond the zone of protection of the relayed breaker or recloser at the substation, additional overcurrent protective equipment usually will be installed out on the main feeder. Manually operated sectionalizing equipment such as pole-top disconnecting switches or solid blade cutouts also are installed at strategic locations along the main feeder to

Provide a convenient means of isolating faults so that repairs can be made after other parts of the feeder are restored to service

Provide means of connecting the feeder to adjacent feeders so that service can be maintained to most customers while repair or maintenance operations are taking place

On underground feeders, this sectionalizing equipment is often in the form of 3-phase, manually operated, load-break switches.

Branch-Circuit Protection. It is exceedingly important to isolate faults on branch and subbranch lines, even short ones, in order to maintain service on the rest of the feeder. Not only does the branch-circuit protection protect the rest of the feeder, but it helps to pinpoint the location of the fault.

Also, there is usually much more mileage and much more exposure in the branch circuit or laterals than in the feeder main. The simple expulsion-fuse cutout is almost universally used for branch and subbranch overcurrent protection. It may be used in combination with reclosers.

On underground feeders, the lateral circuits usually are fused at the point where the main feeder is tapped to establish the lateral. Often, the fuses for several lateral circuits are grouped into a sectionalizing equipment which may also incorporate main-feeder and load-break sectionalizing switches.

Temporary Fault Protection. On overhead distribution circuits, a large portion of the faults are of a temporary nature or are potentially of a temporary nature. For example, some types of transitory faults include momentary contacts with tree limbs and lightning flashover of insulators or crossarms where no sustained 60-Hz short-circuit current is established and no protective devices operate. Other types of faults which result in 60-Hz follow current can be of a transient nature if the circuit voltage can be removed quickly for a short period of time and then restored after the fault path has recovered adequate dielectric strength. Such faults can result from lightning flashovers, bird or animal contacts, conductors swinging together, etc. Reclosers and reclosing breakers provide the function of fault deenergization, pause for deionization of the arc path, and reestablishment of voltage.

If the fault has disappeared during the "dead time," the reclosure is successful. If not, one or more additional reclosing cycles may be attempted. If the fault persists after the prescribed number of reclosing operations, the breaker or recloser will lock open, or the fault will be removed by operation of a fuse or sectionalizer.

It should be recognized that the reclosing function is provided to eliminate the effects of *temporary* faults only. If all faults were of a permanent nature, reclosing would be pointless. Also, temporary faults on branch circuits result in a momentary outage to all customers on the feeder when reclosing is used. Some utilities, in an effort to reduce the number of momentaries, are allowing the branch fuse to blow for temporary faults. (This is done by eliminating the instantaneous trip.) While this procedure reduces the number of momentaries seen by customers, it has the negative effect of creating a substantial interruption out of a temporary fault condition for the customers on the affected branch.

To provide effective protection against temporary faults, all parts of the feeder should be within the zone of a reclosing device. That is, if the station recloser or relayed circuit-breaker sensing does not reach to the remote ends of the circuit, it should be supplemented with reclosers out on the line. (The term *reach* here is used with the meaning of "sense" faults or "sense and operate" for faults.)

Permanent Fault Protection. Permanent faults are those which require repairs, maintenance, or replacement of equipment by the utility operating department before voltage can be restored at the point of fault. System overcurrent protection is provided to disconnect the faulted portion of the system automatically so that an outage is experienced by a minimum number of consumers. Isolation of permanent faults is usually accomplished by the operation of fuse cutouts. It is also achieved in some cases by operation (to lock out) of reclosers, circuit breakers, or sectionalizers.

Combination of Permanent and Temporary Fault Protection. If all faults were of a permanent nature, low-cost fuse cutouts would be the best solution for primary line protection. If all faults were temporary, automatic reclosing devices capable of covering the entire circuit would be the best solution. In actual practice, both kinds of faults occur, and the problem becomes one of selecting the type of device or combination of devices to provide best overall results. For selection of a system of overcurrent protection, it is necessary to give proper consideration to many factors such as importance of service, total number of faults per year, ratio of temporary to permanent faults, cost to utility of service interruptions, and annual charge on investment.

Selection of Overcurrent Protective Equipment—General. The one-line diagram of a distribution circuit, as shown in Fig. 18-18, will show how a well-coordinated installation of overcurrent protective equipment can be made.

At the left is the substation, which steps down the voltage from high-voltage subtransmission level to primary-distribution voltage level. It is at this point that the distribution system starts. A distribution substation usually has a number of radial 3-phase feeders radiating from it. However, for the purposes of illustration, only a single feeder will be considered, and it is shown extending to the right from the substation. At various points along the feeder, branch lines or laterals are tapped off and in some cases subbranches are tapped from these branches. There are, of course, loads (residences, stores, garages, etc.) all along the feeder, branches, and subbranches. Only a few of these loads are shown, for the sake of clarity of the diagram.

It is general practice to install a fuse on the primary (incoming) line side of each distribution transformer, as shown in Fig. 18-18. This may be a transformer internal fuse or an external fuse installed in a cutout. Transformer fusing will be discussed later. Figure 18-18 shows the basic system to which additional overcurrent protective equipment must be added to assure good service continuity.

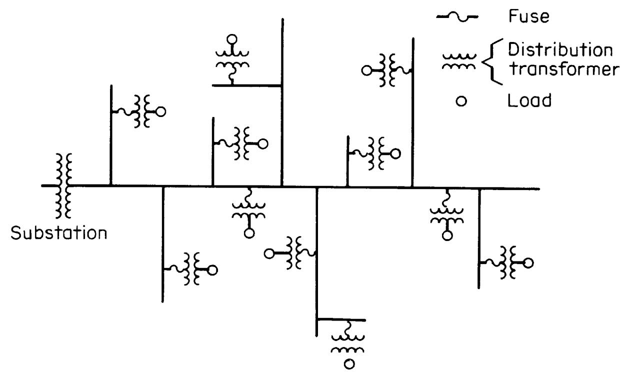


FIGURE 18-18 Distribution feeder.

To properly apply overcurrent protective equipment to this system, it will be necessary to know the highest and lowest (maximum 3-phase and minimum line-to-ground or line-to-line) values of short-circuit currents which can flow if a fault should occur where the feeder leaves the substation, at each branch junction point, and at each subbranch junction point, as well as the minimum line-to-ground short-circuit current which could flow if a fault should occur at the end of any of the branches or subbranches. These short-circuit currents may be calculated easily by conventional methods.

Clearing Nonpersistent or Temporary Faults. Operating records, as well as numerous studies, indicate that a reduction of 75% to 90% in the number of total outages on an overhead system can be attained by the installation of automatic reclosing devices (automatic circuit recloser or reclosing circuit breaker). The recloser or breaker will open the circuit "instantaneously" when a fault occurs, and reclose it after a short period of time.

Referring to Fig. 18-18, automatic circuit reclosers will be applied to protect the entire system against temporary faults. To achieve this sort of protection, the first recloser should be installed on the main feeder at the substation or the power circuit breaker at the substation should be equipped with overcurrent and reclosing relays.

In applying reclosers to do this job, certain factors must be considered: (1) The voltage rating of the recloser must be high enough to meet the requirements of the system. (2) Load current, or the amount of current which flows at the point of installation of the recloser under full-load conditions, should not exceed the amount of current which the manufacturer has rated the recloser to carry continuously (continuous-current rating). Recloser ratings are usually selected to be 140% of the peak load current of the circuit. This allows for normal load growth. (3) The highest value of short-circuit current which will flow through the recloser and which the recloser must interrupt. This value should not be greater than the highest value of current which the recloser is rated to interrupt (interrupting rating). Typically, a recloser will have a continuous rating of 560 A or less and an interrrupting rating of 16,000 A or less. A breaker, on the other hand, will usually handle at least 1200 A continuously and up to about 40 kA under short-circuit conditions.

Referring to Fig. 18-19, a recloser or breaker with reclosing relays will be located at *A* to meet the three application principles mentioned above. This device will be depended on to clear nonpersistent faults which occur in the feeder, branches, or subbranches, anywhere within its protective orbit zone *A* (shown by dotted line in Fig. 18-19). This protective zone extends to the point where the minimum available short-circuit current, as determined by calculation, is equal to the smallest value of current which will cause the device to operate. This value of current required to operate the recloser or breaker is called minimum pickup current. For a recloser it is usually equal to *twice the continuous current rating of the recloser*. A fault beyond this zone may not cause the recloser or breaker *A* to operate, and therefore, another recloser, *B*, with a lower minimum pickup current rating, should be installed just inside of zone *A*, thus resulting in so-called overlapping protection.

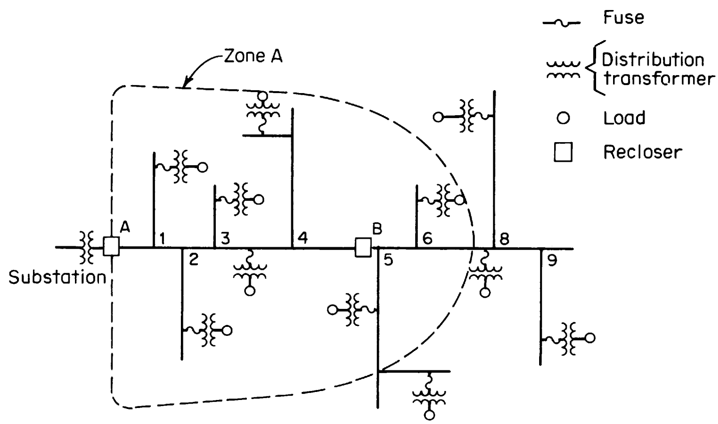


FIGURE 18-19 Distribution feeder with automatic reclosers.

This second recloser, *B* in Fig. 18-19, is placed on the source side (side nearest source of power) of branch 5 so that it can protect the end of this branch from nonpersistent faults which may not cause recloser *A* to operate. It is applied according to the same considerations as was the recloser at *A*. It will be assumed that a fault on the feeder or any branch or subbranch beyond (to the right of) *B* will cause enough current to flow to operate the recloser at *B*. Every point on the entire circuit is now protected against nonpersistent faults because every point is within the protective zone of some reclosing device. Obviously, if every point were not within the protective orbit of some reclosing device, another recloser would have to be installed still farther out on the line.

Clearing Persistent Faults. The first requirement of protecting the circuit against nonpersistent or transient faults has been taken care of by recloser application. It is necessary now to concentrate on the second and third requirements, that is, confining persistent faults to the shortest practical section of line and making persistent faults easy to locate.

If a permanent fault occurs anywhere on the system beyond a recloser, the recloser will operate once, twice, or three times instantaneously, depending on adjustment, in an attempt to clear the fault. However, since a persistent fault will still be on the line at the end of these operations, it must be cleared by some means other than the instantaneous recloser operations. For this reason, the recloser is provided with one, two, or three time-delay operations, depending on adjustment. These additional operations are purposely slower (time-delay operations) to provide coordination with fuses or to allow the fault to "self-clear." If the fault is still on the line after the last opening, the recloser will not close in but lock open.

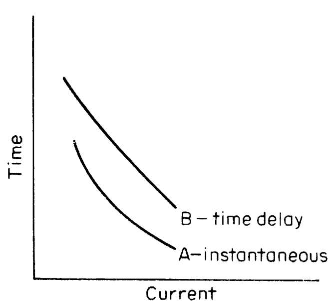


FIGURE 18-20 Recloser tripping characteristics.

Referring to Fig. 18-20, curve *A* represents the instantaneous tripping characteristic with respect to time for the first and second opening of a conventional automatic circuit recloser. Curve *B* represents the tripping characteristics for the third and fourth openings. Following the fourth trip on time delay, the recloser will lock out and must be manually reclosed after the cause of the fault has been remedied.

A persistent fault on a branch or subbranch line *should not* cause a recloser to lock open, since a fault on a relatively unimportant subbranch could shut down the entire circuit, in addition to being extremely difficult to locate. Therefore, some means should be employed to confine outages due to persistent faults to the branch or subbranch on which they occur. This may be done in either of two ways.

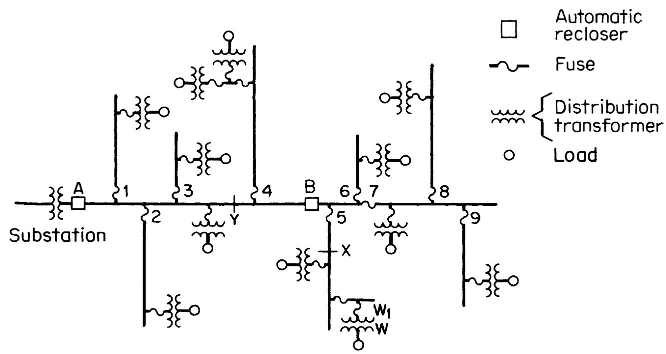


FIGURE 18-21 Distribution feeder with automatic reclosers and fuse cutouts.

One method by which persistent faults can effectively be dealt with is illustrated in Fig. 18-21. A fuse cutout is installed at each branch or subbranch junction to confine outages due to persistent faults to the branch or subbranch on which they occur, that is, fuses 1, 2, 3, 4, etc.

The fuse cutout to be installed at a particular location must be of sufficiently high voltage rating to meet the voltage requirements of the circuit. Its continuous current rating must be equal to or greater than the full-load current at the point of installation. Its interrupting rating must be high enough so that it will successfully open the circuit for any persistent fault occurring beyond it. This may be checked by comparing the interrupting rating of the cutout with the maximum available short-circuit current calculated for the point on the system where the cutout is to be installed.

For an ideal system, when the correct ratings of fuse links are used throughout the system, *no fuse will be blown or even damaged by a temporary fault beyond it; that is, the recloser will open the circuit one, two, or three times on instantaneous operations without the fuse link being damaged*. In many systems, however, where short-circuit levels are very high, it is sometimes impossible to prevent even the largest fuse from operating during a temporary fault. On a permanent fault, the first fuse link on the source side of the fault will be blown, and the circuit thus will be opened by the blowing of the fuse during the third or fourth (time-delay) operation of the recloser, before the recloser will lock open. Hence, the fault will be isolated by the fuse, and the recloser will reset automatically, restoring service everywhere except beyond the blown fuse. The recloser should never lock open on a permanent fault beyond the fuse if it has been properly coordinated with the recloser. Extensive coordination tables are available, as illustrated in Table 18-4, to simplify and facilitate the job of coordinating reclosers with fuse links.

Recloser-Fuse Coordination. Figure 18-22 shows the time-current characteristic curves of the automatic circuit recloser similar to those shown in Fig. 18-20. On these curves, the time-current (TC) characteristics of a fuse *C* are superimposed. It will be noted that fuse curve *C* is made up of two parts; that is, the upper portion of the curve (low current range) represents the total clearing-time TC curve, and the lower portion (high current range) represents the melting TC curve for the fuse. The intersection points of the fuse curves *C* with the recloser curves *A* and *B* illustrate the limits between which coordination will be expected. Basically, this is correct within the interest of simplicity. However, to establish intersection points *a* and *b* accurately and to prepare coordination charts, it is necessary that the characteristic curves of both recloser and fuse be shifted, or modified, to take into account alternate heating and cooling of the fusible element as the recloser goes through its sequence of operations. For example, if the fuse is to be protected for two instantaneous openings, it is necessary to compute the heat input to the fuse during these two instantaneous recloser operations.

Curve A' in Fig. 18-23 is the equivalent TC characteristic of two instantaneous openings (*A*) and is compared with the fuse-damage curve, which is 75% of the melting-time curve of the fuse. This will establish the high current limit of satisfactory coordination indicated by intersection point b'. To establish the low current limit of successful coordination, compare the total heat input to the fuse

TABLE 18-4 Automatic Recloser and Fuse Range of Coordination*

Recloser rating, rms A (continuous)		Fuse link ratings, rms A							
		25T	30T	40T	50T	65T	80T	100T	140T
		Range of coordination, rms A							
50	Min	190	480	830	1200	1730	2380		
	Max	620	860	1145	1510	2000	2525		
70	Min	140	180	365	910	1400	2000	2750	
	Max	550	775	1055	1400	1850	2400	3200	
100	Min	200	200	200	415	940	1550	2280	
	Max	445	675	950	1300	1700	2225	3050	
140	Min		280	280	280	720	710	1750	
	Max		485	810	1150	1565	2075	2875	
200	Min				400	400	400	880	3200
	Max				960	1380	1850	2600	4000
280	Min						620	620	1350
	Max						1500	2200	4000

*Recloser sequence: two instantaneous plus two standard time-delay operations.

represented by curve B', which is equal to the sum of two instantaneous (A) plus two time-delay (B) operations, with the total clearing-time curve of the fuse. The point of intersection is indicated by a'.

On the basis of all corrections added, the fuse will coordinate successfully with recloser between the current limits of a' and b'.

To further clarify what is meant by coordination within prescribed limits, refer to Fig. 18-21—branch 5 and recloser *B*—and also Fig. 18-23 to establish how coordination is achieved between the limits of a' and b'. Assume that fuse 5 beyond recloser *B* is to be protected against blowing or being damaged during two instantaneous operations of the recloser *in the event of a transient fault at X*. If the maximum calculated short-circuit current at the fuse location does not exceed the magnitude of current indicated by b', the fuse will be protected against blowing during all transient faults. By observation of the characteristics in Fig. 18-23, for any magnitude of short-circuit current less than b' but greater than a', the recloser will trip on its instantaneous characteristic once or twice to clear the fault before the fuse-melting characteristic is approached. On the other hand, *if the fault at X is persistent*, the fuse at 5 should blow before the recloser *B* locks out. If the minimum (line-to-ground) calculated short-circuit current available *at the end of branch* 5 is substantially greater than the

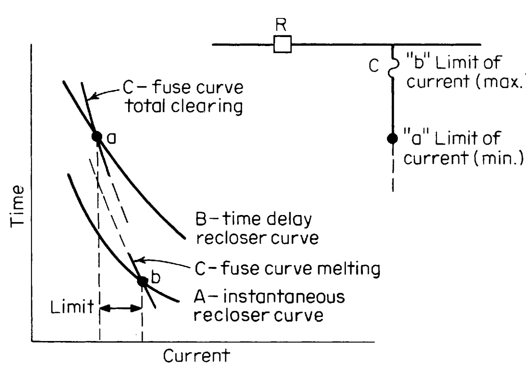


FIGURE 18-22 Recloser and fuse time-current characteristics.

current indicated by a', the fuse will blow (Fig. 18-23) in accordance with the total clearing characteristic, probably before the first time-delay characteristic of the recloser is approached.

The correct fuse link for any application may be selected by comparing its TC characteristics curve with those of the recloser and making certain allowances and corrections as shown. However, tables have been prepared similar to Table 18-4 to simplify greatly the job of coordinating reclosers with fuse links. This table shows the maximum and minimum currents at which certain ratings of fuse links will coordinate with certain ratings of reclosers. The only requirement in their use is a knowledge of the available short-circuit currents and load currents on the system.

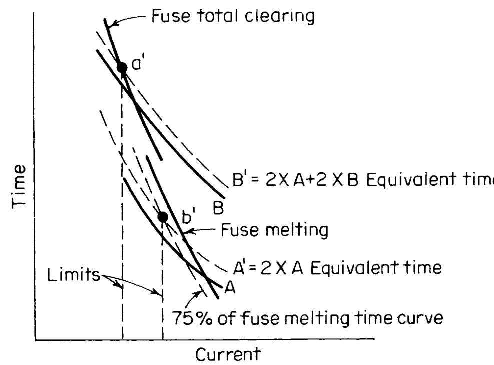


FIGURE 18-23 Recloser and fuse time-current characteristics.

Other sequences of recloser operation can be employed, but one instantaneous and two time-delay operations is the combination most widely used. In some cases, it is necessary to coordinate recloser operation with a relayed breaker at the substation. The principles of coordination are similar to the previous discussions, but a detailed study is beyond the scope of this handbook. This is also true of the application requirements for power-class reclosers for substation and line protection.

Fuse-to-Fuse Coordination. It may be desirable to use more than two fuses in series beyond a recloser in order to reduce the number of consumers affected by an outage. An example of this would be the fuses at points 7, 8 and at transformers on branch 8 in Fig. 18-21. The coordination of these fuses in series beyond the recloser *B* may be accomplished by coordinating adjacent fuses first with each other and then with the recloser in the manner just outlined.

Figure 18-24 illustrates the general principle of coordinating fuses in series. Fuse 7 is called the *protected fuse*, and fuse 8 is called the *protecting fuse*. For perfect coordination, fuse 8 must clear the circuit during a fault anywhere beyond it, such as at *X*, before fuse 7 is damaged or partially melted. From this can be seen the requirement for melting-time–current curves plotted to minimum values and total-clearing-time–current curves plotted to maximum values for each fuse-link rating. *Total-clearing-time* curves represent the total time, including melting time and arcing time, plus manufacturing tolerance, that it takes the fusible elements to clear the circuit. *Melting-time* curves represent the minimum time, based on factory test, at which the fusible element melts for various currents. From the melting-time curves, *damaging-time* curves can be determined by applying a factor of safety. It usually is suggested that the damaging-time curve be made by taking 75% of the melting time (in seconds) of a particular size at various current values.

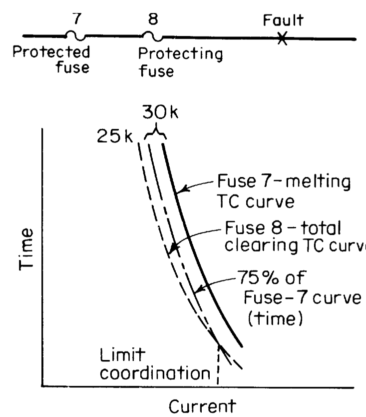


FIGURE 18-24 Fuse time-current characteristics.

To establish coordination of two fuses in series, it is necessary to compare the total-clearing-time–current curve of the protecting fuse with the damage-time–current curve of the protected fuse. If there is no intersection of these two curves throughout their entire current range, coordination or selectivity can be expected. Where there is an intersection of the curves, the current value indicated by the point of intersection will establish the limit of selectivity.

TABLE 18-5 Fuse Ratings

Type K EEI-NEMA ratings, A, of the protecting fuse links (8 in diagram)	Type K EEI-NEMA ratings, A, of the protected fuse links (7 in diagram)								
	6K	8K	10K	12K	15K	20K	25K	30K	40K
	Max short-circuit rms A to which fuse links will be protected								
1K	135	215	300	395	530	660	820	1100	1370
2K	110	195	300	395	530	660	820	1100	1370
3K	80	165	290	395	530	660	820	1100	1370
5-A series Hi-surge	14	133	270	395	530	660	820	1100	1370
6K		37	145	270	460	620	820	1100	1370
8K			133	170	390	560	820	1100	1370
10-A series Hi-surge		16	24	260	530	660	820	1100	1370
10K				38	285	470	720	1100	1370
12K					140	360	660	1100	1370
15K						95	410	960	1370
20K							70	700	1200
25K								140	580

Source: General Electric Company.

Because of the inherent characteristics of fuses, the maximum available short-circuit current in that section (determined by calculation) controlled by the protecting link (8 in Fig. 18-24) is the determining current which establishes coordination possibilities.

Most fuse-link manufacturers publish tables which make coordination very simple. These tables eliminate the necessity of comparing actual fuse-characteristic curves. Table 18-5 is illustrative of tables used for fuse-to-fuse coordination. The values in the left-hand column are the protecting fuse ratings and the values across the top are the protected fuse ratings. The numerical values in the table show the magnitude of current or curve intersection points at which, or below which, fuse 7 will be protected by fuse 8. These current magnitudes are maximum values; in other words, for any short-circuit current greater than that shown, fuse 7 will be damaged. Hence, a larger-rated fuse will have to be selected for location 7 or else its position must be changed.

Isolation by Sectionalizer. Another method of isolating persistent faults is to install a device, known as a *sectionalizer*, at locations where a fuse might otherwise be used. A sectionalizer is a device which counts the operations of a backup automatic-interrupting device such as a recloser. It has no interrupting capacity of its own but operates in a predetermined coordination scheme to open a faulted lateral before the backup device locks out.

The sectionalizer opens the circuit after a predetermined number (usually two or three) of operations of a reclosing device. Its opening operation occurs during a period when the reclosing device is open. It can be used to replace a lateral sectionalizing fuse or to replace a lateral recloser where interrupting requirements have grown beyond the capability of the recloser. Among its operating advantages are

It allows coordination with breakers or reclosers where fault current is above 5000 A. Such coordination usually is impossible with expulsion fuses.

It can provide a new sectionalizing point on an existing circuit without upsetting existing overcurrent coordination, since the device operates as a counter and does not introduce another level of time-current coordination.

Equipment Protection

General. It is necessary to provide overcurrent protection for distribution equipment such as capacitors and distribution transformers:

To protect the system from the effects of equipment failures

To reduce the probability of violent failures

To indicate the location of the fault

A detailed discussion of all aspects of overcurrent protection of equipment is beyond the scope of this handbook. However, because of its importance, a few comments will be included regarding the overcurrent protection of distribution transformers.

Self-Protected Transformers. The term *self-protected distribution transformer* is applied to units which incorporate an internal primary expulsion fuse, a direct-mounted arrester, and an internal secondary circuit breaker. The low-voltage circuit breaker protects the transformer from excessive overload and from some of the faults originating on the secondary system. The expulsion fuse has the sole function of removing a failed transformer from the system.

The rating of the internal expulsion fuse usually is quite large compared with the continuous current rating of the transformer, perhaps 10 to 14 times. This is done

1. To ensure that the fuse is not damaged by the maximum tripping current of the circuit breaker
2. To minimize the possibility of extraneous fuse blowing because of lightning current effects

Another reason is that fuse removal and replacement may require that the transformer be taken to a shop facility.

Transformer internal expulsion fuses are installed at the factory and are given a designating number rather than an ampere rating for coordination purposes. For a 7200-V transformer, the internal expulsion fuse, often called *weak link*, has an interrupting capacity of about 3000 A. Weak links for higher-voltage transformers have somewhat lower interrupting capacity.

Despite the fact that self-protected transformers often are installed at locations on the system where the interrupting capacity of the weak link may be exceeded for a solid fault, experience over the years has been excellent, probably because most transformer failures begin as relatively low fault-current turn-to-turn failures. As the fault current progressively becomes larger, the fuse will operate well before its interrupting capacity is exceeded. Thus, while high-current transformer faults can occur, their frequency of occurrence is very small.

However, there is growing concern among utility companies regarding the occasional violent failures of transformers, and many users are using, or are considering the use of, current-limiting fuses as one method to minimize the energy input into a failed transformer.

The secondary circuit breaker is depended on to provide protection against excessive transformer loads and secondary system faults that occur within its zone of protection, or *reach*. Its TC characteristic should be such that it will always operate before the primary fuse suffers any damage, as illustrated in Fig. 18-25. On the other hand, the breaker should not operate for faults beyond the customer's service-entrance-protective equipment. Likewise, the internal primary fuse should operate to clear transformer faults before damage occurs to the line sectionalizing fuses back toward the source.

Conventional Transformers. Conventional distribution transformers usually are protected by separately mounted expulsion fuse cutouts in series with the primary winding. No secondary overcurrent protection is provided, so protection against extreme overloads or secondary faults, if any, must come from the primary fuse. Therefore, the size of the primary fuse is relatively much smaller than for the self-protected transformer, usually being chosen in the range of 2 to 3 times the full-load current of the transformer.

It is desirable to keep the fuse rating as low as possible consistent with certain application limitations:

1. When a transformer is energized by closing of its cutout or operation of a recloser or other switch, a large "magnetizing inrush" current can occur. Initially, this current can be as much as 20 or more times normal, rapidly decaying to normal in a short time—perhaps $^1/_2$ to 1 s or more. The primary fuse link must be large enough to avoid damage by the magnetizing inrush current, so it usually is selected at least large enough to carry 12 times rated transformer current for 0.1 s without damage.
2. The primary fuse should not be damaged by lightning currents or arrester discharge currents (depending on connection used) or large magnetizing currents which can result from saturation of the core due to lightning currents. Many utilities assign an arbitrary minimum fuse size which they will employ. With expulsion fuses, 10- or 15-A rating is often designated as the minimum size.

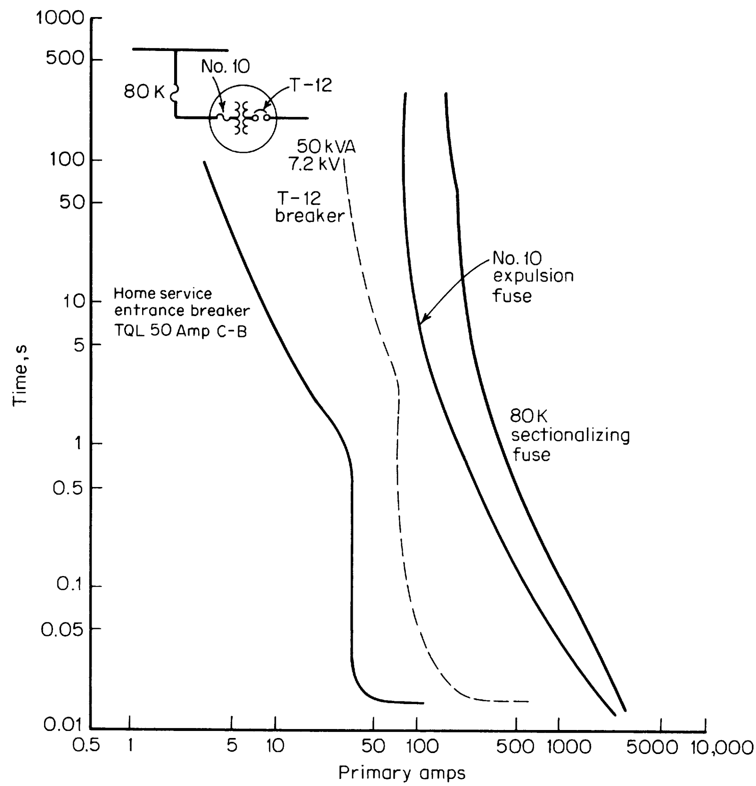


FIGURE 18-25 Overcurrent coordination for self-protected distribution transformer.

With a fuse *rating* of 2 to 3 times rated transformer current, the minimum melting current under long-time conditions will be in the range of 4 to 6 times transformer rating. Consequently, little overload protection is obtained.

In the absence of overload protection, many utilities count on a transformer load-management program or seasonal load-survey techniques to keep their "burnouts" at an acceptable level. Also, the primary fuse has a limited reach as far as secondary faults are concerned; therefore, secondary faults can occur which cannot be "seen" by the fuse. Often these faults—especially on underground systems—will burn clear.

Expulsion Cutouts. Distribution expulsion cutouts are by far the most common type of protective device used on overhead primary-distribution systems. The open-type cutout has generally supplanted the porcelain-enclosed style. The cutout consists of an insulating structure and a hinged fuse tube of hollow cylindrical construction which contains the fuse link. When the fuse link melts, the ensuing arc impinges on the wall of the fibrous tube holder (and usually a small auxiliary tube), generating gas which provides the expulsion action needed to extinguish the fault current. Separation of the fuse link also releases the cutout-latching mechanism so that the fuse holder falls to the open position and can readily be located by operating personnel. The fuse holder also can be switched manually with a switch stick, much like a disconnect switch. In some cases, a solid blade is used in place of the fuse holder to provide a disconnecting function. The cutout also can be provided with load-breaking accessories so that it can be used as a load-break switch.

Generally cutouts are available in 100- and 200-A continuous-current ratings for fuses and 300- or even 400-A with solid blades. Cutouts are available with voltage ratings for all the common primary system voltages and interrupting capacities generally from 1200 up through 16,000 A symmetrical and more.

Fuse links for cutouts are available with a variety of TC characteristics. However, the two most widely used types are the Type K (fast) fuse links and the Type T (slow) links with characteristics as defined in ANSI C37.43. Both types have certain application advantages and disadvantages which must be evaluated by the utility. Ordinarily, a given utility uses one type or the other, not both. Use of the Type K links is believed to be somewhat greater than use of Type T.

Other common types of primary fuses employ fusible elements immersed in oil.

Fuses are not widely used in electric utility secondary systems, with the notable exception of secondary network systems, where *limiters* are frequently used in the secondary cable circuits. Limiters are fusible elements whose TC characteristics are coordinated with the cable size and insulation characteristics to prevent damage to the cable when faults do not burn clear or self-extinguish.

Current-Limiting Fuses. The use of current-limiting fuses in distribution systems has been growing. The fuse generally is constructed of silver wire or ribbon fusible elements—often several in parallel—spirally wound on a core or spider and packed in a quartz-sand filler in a sealed cylindrical glass or epoxy-glass container. Provisions for suitable electrical connections are made at the ends. When operation takes place under high-fault-current conditions, the fusible element melts almost instantaneously at a series of reduced sections all along its length. The resulting arc dissipates its heat rapidly into the surrounding sand, melting the sand around the arc into a glass-like structure called a *fulgorite*. This action builds up the apparent resistance of the fuse extremely rapidly, resulting in a "back voltage" greater than system voltage. Thus, the fault current is limited to a value much less than the available system fault currrent.

Current-limiting fuses are characterized by

1. High-current interrupting ability. Interrupting ratings of 50,000 A symmetrical or greater are commonly available.
2. Operation is noiseless, and there is no expulsion of the arc or arc products. Thus, the fuse can be "packaged" into relatively confined space in transformers and protective equipment, making it extremely attractive for use on underground systems.
3. In the current-limiting mode of operation, the interrupting time is very fast, one-half cycle or less.
4. Current-limiting action and fast operation reduce the amount of I^2t (or fault energy) let through into failed equipment, thereby reducing resultant damage. In the case of distribution transformers applied on systems of high available fault current, protection by current-limiting fuses can virtually eliminate violent failures due to high fault current.

General-purpose current-limiting fuses are designed to clear fault currents over a broad range. They are defined by ANSI Standard C37.40-3.2.2.2 as fuses capable of interrupting all currents from the maximum interrupting current down to the current causing melting of the fusible element in 1 h. Current-limiting fuses inherently are excellent fault-current interrupters in the high current range. Typical general-purpose fuses operate in the current-limiting mode at fault currents equal to *approximately* 25 times rated current or larger.

Special design and construction techniques are required to obtain clearing of low-fault-current values. For operating times greater than about 0.01 s, the fuses have TC characteristics which are plotted on loglog coordination paper in the same manner as expulsion fuse characteristics.

Backup current-limiting fuses are defined by ANSI Standard 37.40-3.2.2.1 as fuses capable of interrupting all currents from the rated maximum interrupting current down to the rated minimum interrupting current as given by the manufacturer. The low current clearing must be accomplished by an auxiliary device, most commonly an expulsion fuse. In this case, the TC characteristics are a composite of the two fuses as shown in Fig. 18-26. The backup current-limiting fuse can be retrofitted into existing pole-type distribution transformer installations which have expulsion fuse protection only.

18.10 OVERVOLTAGE PROTECTION

Lightning. Lightning is the most frequent cause of overvoltages on distribution systems. Basically, lightning is a gigantic spark resulting from the development of millions of volts between clouds or between a cloud and the earth. It is akin to the dielectric breakdown of a huge capacitor.

The voltage of a lightning stroke may start at hundreds of millions of volts between the cloud and earth. Although these values do not reach the earth, millions of volts can be delivered to the building, tree, or distribution line struck. In the case of overhead distribution lines, it is not necessary that a stroke contact the line to produce overvoltages dangerous to equipment. This is so because "induced voltages" caused by the collapse of the electrostatic field with a nearby stroke may reach values as high as 300 kV.

The amount of current in a stroke is a statistical quantity, depending on the energy in the cloud and the voltage difference between the cloud and the earth at the start of the stroke. A few stroke currents in excess of 200,000 A have been measured; however, 50% of all stroke currents are less than 15,000 A.

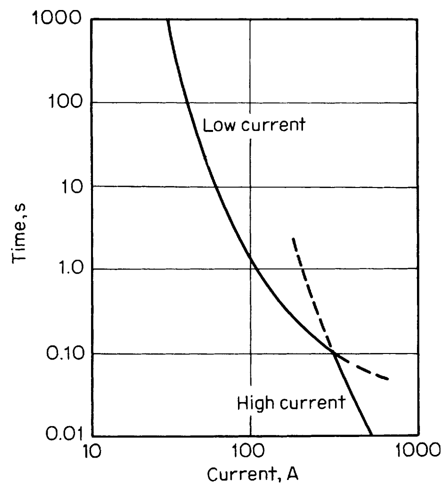


FIGURE 18-26 Time-current characteristic for backup current-limiting fuse in series with expulsion fuse.

The time duration of the current flow in the majority of the high-current strokes is only tens or hundreds of microseconds. Typically, the current rises to its maximum in 0.5 to 10 μs, decreases to half value in 20 to 50 μs, and falls to zero within 100 to 200 μs. On a 60-Hz basis, these are extremely short times if one considers that one-half cycle is equivalent to 1/120 s or 1,000,000/120 = 8333 μs. Numerous field investigations have established the numerical statistics which apply to lightning.

In summary, lightning can produce voltages dangerous to the distribution system and all its component equipment. It poses a major threat to service continuity and must be coped with by means of distribution surge arresters.

Arrester Selection. Choosing an arrester rating for a distribution system is based on the system's line-to-ground voltage and the way it is grounded. The limiting condition for an arrester does not usually have anything to do with the magnitude of the surges (switching or lightning) that it might see. This is in contrast to the selection of arresters for transmission. In distribution, rating of the arrester is based on the maximum steady-state line-to-ground voltage the arrester might see. This limiting condition is normally caused when there ia a line-to-ground fault on one of the other phases.

According to ANSI Standard C62.22, "Guide for the Application of Metal-Oxide Surge Arresters for Alternating-Current Systems," proper application of arresters on distribution systems requires knowledge of "(1) the maximum normal operating voltage of the power system, and (2) the magnitude and duration of temporary overvoltages (TOV) during abnormal operating conditions. This information must be compared to the arrester MCOV rating and to the arrester TOV capability."

The MCOV of the arrester is, however, somewhat easier to define because it is approximately 84% of the arrester duty cycle rating. What this means is that a 10-kV duty cycle rated arrester, typically used for a 13.2-kV system, could be operated continuously with a maximum continuous line-to-ground voltage of 8.4 kV or less.

TABLE 18-6 Commonly Applied Voltage Ratings of Metal-Oxide Arresters on Distribution Systems[a]

System voltage, kV rms		Commonly applied arrester voltage ratings, kV rms duty cycle voltage ratings (MCOV)[b]		
Nominal voltage	Maximum voltage range B[c]	4-Wire multigrounded neutral wye	3-Wire low-impedance[d] grounded[e]	3-Wire high impedance[d] grounded
2400	2540			3 (2.55)
4160Y/2400	4400Y/2540	3 (2.55)	6 (5.1)	6 (5.1)
4260	4400			6 (5.1)
4800	5080			6 (5.1)
6900	7260			9 (7.65)
8320Y/4800	8800Y/5080	6 (5.1)	9 (7.65)	
12000Y/6930	12700Y/7330	9 (7.65)	12 (10.2)[f]	
12470Y/7200	13200Y/7620	9 (7.65) or 10 (8.4)	15 (12.7)[f]	
13200Y/7620	13970Y/8070	10 (8.4)	15 (12.7)[f]	
13800Y/7970	14605Y/8430	12 (10.1)	15 (12.7)[f]	
13800	14520			18 (15.3)
20780Y/12000	22000Y/12700	15 (12.7)	21 (17.0)[f]	
22860Y/13200	24200Y/13870	18 (15.3)	24 (19.5)[f]	
23000	24340			30 (24.4)
24940Y/14400	26400Y/15240	18 (15.3)	27 (22.0)[f]	
27600Y/15930	29255Y/16890	21 (17.0)	30 (24.4)[f]	
34500Y/19920	36510Y/21080	27 (22.0)	36 (29.0)[f]	

[a]Spacer cable circuits have not been included—there has been insufficient experience with the application of metal-oxide arresters on spacer cable circuits to include them in this table.
[b]For each duty cycle rating, the maximum continuous operating voltage (MCOV) is also listed.
[c]See ANSI C84.1-1989.
[d]Low impedance circuits are typically 3-wire, unigrounded at the source. High impedance circuits are generally ungrounded (i.e., delta). Additional information regarding system grounding is contained in ANSI C62.92, Part 1.
[e]Line-to-ground fault duration not to exceed 30 minutes. For longer durations consult manufacturers' temporary overvoltage capability.
[f]Individual case studies may show lower voltage ratings may be used.

Table 18-6, from ANSI C62.22, shows the commonly applied voltage ratings of metal-oxide arresters for distribution systems. All these duty cycle ratings are the same as the rating for the older gapped silicon carbide arresters except at the 13.8-kV level. Typically, a 13.8-kV, 4-wire, multigrounded system has used 10-kV gapped arresters. Today, most of these same utilities are still using 10-kV MOVs. Some utilities, however, have recognized that the 10-kV arrester is very marginal and possibly should be replaced by a 12-kV rating to be on the more conservative side.

TOV. How much voltage shift which will occur is a function of the type of system grounding. For example, on a delta system, a line-to-ground fault will cause a full offset; that is, the line-to-ground voltage will become the line-to-line voltage. Figure 18-27 illustrates this condition. As can be seen, when a phase has a fault there is no current because the transformer is delta-connected. For a 4-wire multigrounded system, there is less voltage rise (Fig. 18-28). The arresters connected from nominal line-to-neutral voltage multiplied by the product of the regulation factor 1.05 and the voltage rise factor 1.2. This is equivalent to 1.25 times nominal line-to-neutral system voltage. For an MOV type arrester, this voltage is compared with the TOV rating of the MOV. Because the MOV arrester is more sensitive to poor grounding, poor regulation, and the reduced saturation sometimes found in new transformers, it is generally recommended that a 1.35 factor be considered for MOVs.

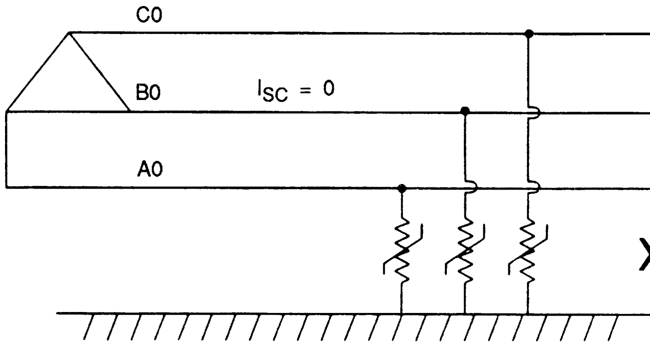


FIGURE 18-27 Line-to-ground fault on a delta system.

A summary of this and other recommendations is as follows:

Open-wire multigrounded system:

$$\text{Rating} = \text{nominal line-to-ground (L-G) voltage} \times 1.25 \text{ (gapped)}$$

$$\text{Rating} = \text{nominal L-G voltage} \times 1.35 \text{ (MOV)}$$

Spacer cable systems:

$$\text{Rating} = \text{nominal L-G voltage} \times 1.5$$

$$\text{Unigrounded rating} = \text{nominal L-G voltage} \times 1.4$$

The temporary overvoltage capability of the arrester as a function of time is shown in Fig. 18-29.

Insulation Coordination

Margins for Overhead Equipment. It is important to note that application of arresters for transmission and distribution is different. In transmission, lightning is of secondary concern in surge

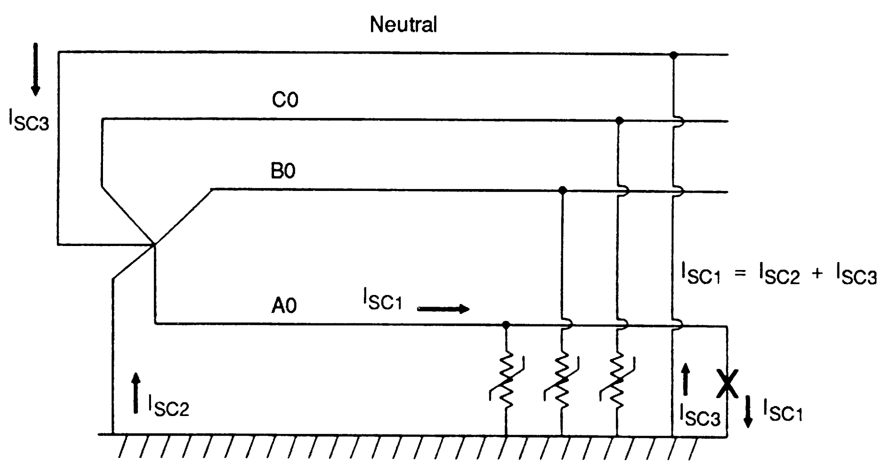


FIGURE 18-28 Line-to-ground fault on a grounded-wye system.

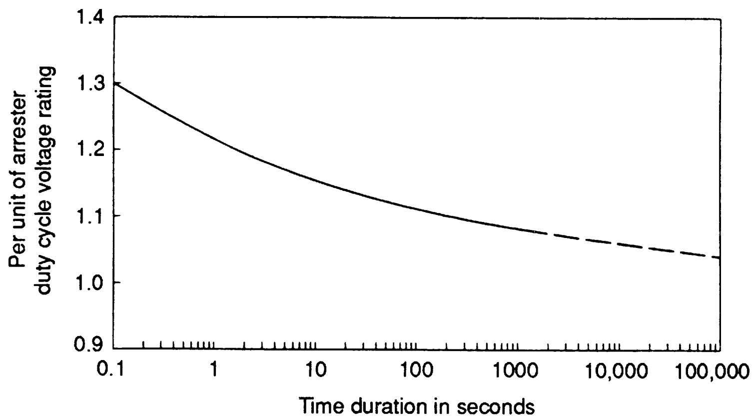


FIGURE 18-29 Example TOV curve (consult arrester manufacturer for exact curve).

arrester application. Primary concern is switching surges. On a distribution circuit, however, the relatively low voltage and short lines tend to make switching surges minimal, and consequently, lightning is of primary importance.

Reflection of this fact can be seen in typical characteristics published for distribution-class arresters, as shown in Tables 18-7 and 18-8. As can be seen, protective characteristics are shown for front-of-wave sparkover and *IR* discharges but not for switching surge waves (as shown for higher rated transmission arresters).

The two protective characteristics normally used for insulation coordination are

Front-of-wave sparkover. This is the first thing that happens to the gapped arrester—it sparks over. It is compared with the fast front equipment insulation characteristics such as the chopped wave insulation level of the transformer. An MOV has no gap but does have an equivalent sparkover, as shown in Table 18-8.

IR discharge at 10 kA. After the arrester sparks over the gap, the lightning current discharges through the block material. Standards recommend that a 10-kA discharge level be used for coordination purposes. (Discharge characteristics across a MOV are very similar to gapped silicon carbide arresters, so the margin calculation is virtually identical.)

TABLE 18-7 Distribution Arrester Characteristics from Handbook—Silicon Carbide

Arrester rating, kV RMS	Maximum ANSI front-of-wave sparkover, kV crest		Maximum discharge voltage, kV crest at indicated 8 × 20-μs impulse current		
	With disconnector	Externally gapped	5000 A	10,000 A	20,000 A
3	14.5	31	11	12	13.5
6	28	51	22	24	27
9	39	64	33	36	40
10	43	64	33	36	40
12	54	77	44	48	54
15	63	91	50	54	61
18	75	105	61	66	74
21	89	—	72	78	88
27	98	—	87	96	107

TABLE 18-8 Distribution Arrester Characteristics from Handbook—MOV (Heavy Duty)

Arrester rating, kV rms	MCOV, kV rms	Front-of-wave protective level,* kV crest	Maximum discharge voltage, 8 × 20-μs current wave 5 kA	10 kA	20 kA
3	2.55	10.7	9.2	10.0	11.3
6	5.10	21.4	18.4	20.0	22.5
9	7.65	32.1	27.5	30.0	33.8
10	8.40	35.3	30.3	33.0	37.2
12	10.2	42.8	36.7	40.0	45.0
15	12.7	53.5	45.9	50.0	56.3
18	15.3	64.2	55.1	60.0	67.6
21	17.0	74.9	64.3	70.0	78.8
24	19.5	84.3	72.3	78.8	88.7
27	22.0	95.2	81.7	89.0	100.2
30	24.4	105.9	90.9	99.0	111.5
36	30.4	124.8	107.0	116.6	131.3

*Based on a 10-kA current impulse that results in a discharge voltage cresting in 0.5 μs.

Distribution equipment is normally defined as being in a voltage class such as 15 or 25 kV. Most utility equipment is operated in the 15-kV class. A distribution transformer in the 015-kV class is defined by the following insulation characteristics:

60-Hz, 1-min withstand = 34 kV

Chopped wave (short-time) = 110 kV at 1.8 μs

Basic insulation level (BIL) = 95 kV

Assuming a 12,470-V, 4-wire system (7200 V L-G), we would select the arrester rating based on the rules developed in the preceding section, i.e., a 9-kV arrester (gapped).

We can see from Tables 18-7 and 18-8 that a 9-kV gapped arrester has a sparkover of 39 kV and an *IR* discharge at 10 kA of 36 kV. This could be plotted with the transformer characteristics as in Fig. 18-30.

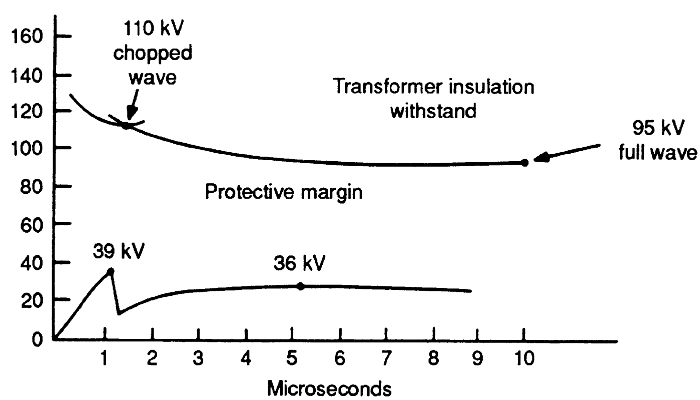


FIGURE 18-30 Insulation coordination.

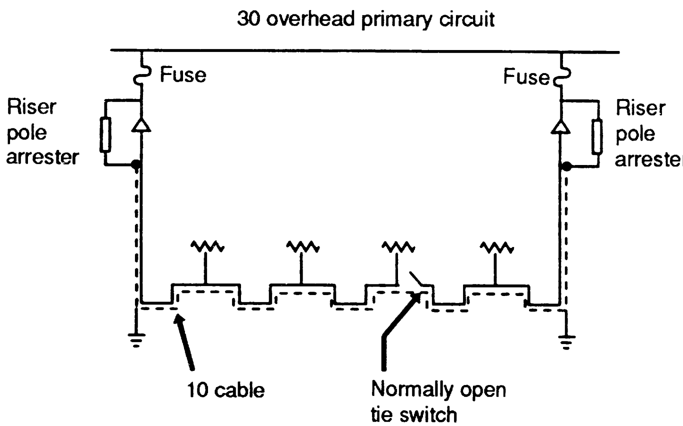


FIGURE 18-31 Underground lateral.

Standards recommend 20% margins calculated by the formula:

$$\text{Margin} = \frac{\text{insulation withstand} - \text{protective level}}{\text{protective level}} \times 100$$

Two margins are calculated, one for the chopped wave and one for the full wave (BIL) of the transformer. These calculations are performed as follows:

$$\text{Margin} = \frac{110 - 39}{39} \times 100\% = 182\% \text{ (chopped wave)}$$

$$\%\ \text{Margin} = \frac{95 - 36}{36} \times 100\% = 164\% \text{ (BIL)}$$

As can be seen, these margins (182% and 164%) are greatly in excess of the recommended 20% and consequently show good protection practice. If we were using an MOV, we would simply use the equivalent sparkover or compare only the *IR* discharge and the BIL, since this is the lesser of the two margins. The margins would be similar.

Margins for Underground Equipment If the system is underground, we must be more concerned with the phenomena of traveling waves and the consequent doubling of voltage surges at an open point. For example, a typical underground residential design is shown in Fig. 18-31. A surge entering the cable will travel to the open point where its voltage will double, as shown in the figure, and start on its way back.

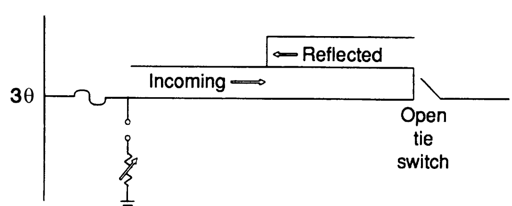


FIGURE 18-32 Reflected surge voltage at open point.

This reflected wave plus the incoming waves impose approximately twice the normal voltage on the entire cable and all the equipment connected to it (Fig. 18-32). For example, if we had an arrester with a 36-kV *IR* discharge level (we are only considering BIL margin), we would now expect to see

72 kV imposed across the insulation of this equipment. The new margin would then be calculated as follows:

$$\% \text{ Margin} = \frac{95 - 72}{72} \times 100 = 32\%$$

For higher voltage levels, where the margins become considerably less, it may be necessary to use an arrester at the open point or riser-pole arresters having lower discharge levels.

18.11 DISTRIBUTION TRANSFORMERS

Distribution transformers convert electrical energy from primary voltages (2.4 to 34.5 kV) to utilization voltages (120 to 600 V). Momentary drops in lighting voltage caused by the starting current of motors often necessitate use of separate transformers where 3-phase motors 20 hp and larger and 1-phase motors 6.5 hp and larger must be served from radial circuits.

Standard Ratings of Single-Phase Distribution Transformers. By agreement between users and manufacturers, certain features of line-transformer design have been standardized for sizes up to 500 kVA and for voltages up to 67,000 V. Capacities are 10, 15, 25, 37 1/2, 50, 75, 100, 167, 250, 333, and 500 kVA.

Voltage ratings on primary windings are 2400/4160Y, 4800/8320Y, 7200/12,470Y, 12,470GrdY/7200, 7620/13,200Y, 13,200GrdY/7620, 12,000, 13,200/22,860GrdY, 13,200, 13,800GrdY/7970, 13,800/23,900GrdY, 13,800, 14,400/24,940GrdY, 16,340, 24,940GrdY/14,400, 19,920/34,500GrdY, 34,500GrdY/19,920, 22,900, and 34,400. On the secondary side, windings are built for 3-wire operation at voltages of 120/240 or for 240/480. For some of the larger kVA sizes, secondary side windings are available at voltages from 2400 to 7970 V. Bushings for secondary terminals are located on the side of the case, except that primary bushings for 7200 V and higher are cover-mounted.

TABLE 18-9 Typical Electrical Characteristics of Single-Phase and 3-Phase 60-Hz Distribution Transformers
Loss factors can vary according to evaluation factors.

Size, kVA	Percent *IR*	Percent *IX*	Percent *IZ*	Percent no-load loss	Percent load loss
Pole-type single-phase transformers—voltage rating 7200/12,470Y to 120/240 V					
10	1.6	1.4	2.1	0.59	1.65
15	1.3	1.0	1.6	0.51	1.28
25	1.2	1.7	2.1	0.38	1.26
37 1/2	1.3	1.9	2.3	0.37	1.31
50	1.1	1.8	2.1	0.36	1.10
75	1.0	2.1	2.3	0.34	1.03
100	1.0	2.1	2.3	0.32	1.02
167	1.0	2.0	2.2	0.29	0.96
250	1.0	2.3	2.5	0.23	0.99
333	0.9	2.4	2.6	0.21	0.90
500	0.8	2.5	2.6	0.20	0.82
Pad-mounted 3-phase transformers—voltage ratings 12,470Y/7200 to 208Y/120 V					
75	1.0	3.0	3.2	0.52	0.95
112.5	1.1	3.2	3.4	0.40	1.15
150	1.0	3.4	3.5	0.39	0.96
225	1.0	3.4	3.5	0.36	0.98
300	1.0	3.8	3.9	0.33	0.97
500	1.0	3.9	4.0	0.27	0.97

Supporting lugs are arranged to permit mounting either by bolting to the pole or by hanging on crossarms. Where necessary, provision is made for a grounding connection to the case or from the secondary neutral terminal to the case. Similar standards have been promulgated by ANSI for 3-phase pole-type transformers up to 500 kVA.

Electrical characteristics typical of single- and 3-phase transformers of the 12470Y/7200V class are given in Table 18-9. Distribution transformers with different primary voltages will have values only slightly different from those shown in Table 18-9. Transformer regulation for a kVA load of power factor cos θ, at rated voltage, can be calculated from the formula

$$\text{Percent regulation} = \frac{\text{kVA load}}{\text{kVA transformer}}\left[\%\,IR\cos\theta + \%\,IX\sin\theta + \frac{(\%\,IX\cos\theta - \%\,IR\sin\theta)^2}{200}\right]$$

Transformers are installed on poles in the following ways: transformers 100 kVA and smaller are bolted directly to the pole, and sizes 167 to 500 kVA have support lugs attached to the transformer and intended for bolting to adapter plates for direct pole mounting or hung on crossarms by means of steel hangers attached securely to the transformer.

Banks of three single-phase transformers are hung side by side on heavy double arms, usually located low on the pole, or on a "cluster" bracket which spaces them around the pole. Three or more transformers 167 kVA and larger are installed on a platform supported by two poles set 10 to 15 ft apart. The transformer-platform structure is often placed on the customer's premises to reduce the distance that secondaries must be run and to avoid pole congestion on public thoroughfares.

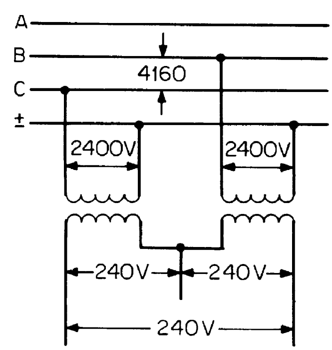


FIGURE 18-33 Open Y connection from 4-wire, 3-phase system.

Transformers are installed in street vaults, in manholes, on pads at ground level, subsurface, or within buildings. When installed within buildings where the possibility of submersion is remote, the overhead or inside types of transformer and cutout are used. Transformer vaults within a building are of fireproof construction, except when transformers are dry type or filled with nonflammable liquid.

Pole-Mounted Regulators. Today 16- or 32-step regulators built for pole mounting cover the customary ±10% voltage range. Voltage-level control and line-drop compensators give them essentially the same characteristics as the larger station regulators.

Open-Δ connection enables small power customers to receive 3-phase service from two transformers connected to a 3-phase circuit, thus reducing the investment in transformers. Open-Δ from a 3-wire system is the usual Δ connection with one transformer omitted. The connection from a 4-wire system is shown in Fig. 18-33, 2-phase wires and neutral being used on the primary side of the transformers. Current in each of two single-phase transformers connected in open Δ is 73% greater than in each of three transformers connected in closed Δ.

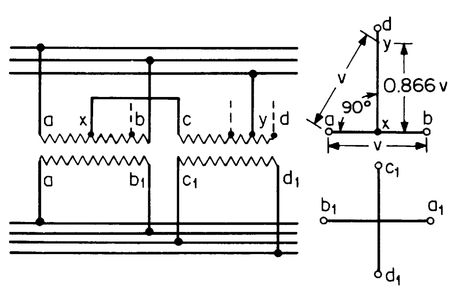


FIGURE 18-34 Balanced T- or Scott-connected transformer connection for 3- to 2-phase transformation.

The *Scott connection* shown in Fig. 18-34 gives an accurate transformation but requires one of the transformers to have an 86.6% tap and the other to have a 50% tap.

18.12 SECONDARY RADIAL DISTRIBUTION

Secondary mains operate at utilization voltage and serve as the local distributing main. In early commercial radial systems, secondary mains that supply general lighting and small power are usually separate from mains that supply 3-phase power because of the dip in voltage caused by starting motors. This dip in voltage, if sufficiently large, causes an objectionable lamp flicker.

Single-phase secondary mains supplying general lighting and small power are usually 3-wire mains operating at 120 V line-to-neutral and 240 V line-to-line. Incandescent lamps, fans, heating devices, small fractional-horsepower motors, and other appliances rated 115 or 120 V are supplied from the 120-V line and neutral. Electric ranges, larger single-phase motors up to 6.5 hp, and large appliances rated 230 or 240 V are supplied 240 V. Some utilities supply these loads at 120/208 V.

Three-phase secondary mains are commonly operated 3-wire 240 V. Some utilities offer 208Y/120-V 3-phase 4-wire service. The 3-phase mains are on the same poles or in the same duct line (but in separate ducts) with single-phase lighting mains. Separate single-phase and 3-phase services are extended to customers who require both types of service. In large commercial and industrial installations, power is often delivered at 480 V to effect an economy in conductor investment.

European practice is to supply 230 V to lighting and appliances from a 230/440-V system. This effects a savings in distribution and interior wiring but results in less efficient incandescent lamps and other small appliances.

In America, large commercial buildings and factories are served at 480Y/277 V because most permanent lighting is fluorescent, which operates efficiently at 277 V, and 480 V is well suited for the numerous 3-phase motors. Such installations have small dry-type transformers to supply 120 V for portable lights, convenience outlets, and tools and for business machines; these transformers are located near the 120-V loads and supplied from the 480-V system.

Fractional-horsepower motors up to about 3/4 hp are regularly supplied by single-phase 120-V mains. Industry committees, sparked by sudden acceptance of home air conditioning, several years ago agreed to permit starting currents not to exceed 50 A for 115-V motors. Special design enabled motors up to 3/4 hp to meet this limitation. Larger motors up to 6.5 hp are usually served at 240 V, although 3- and 6.5-hp motors may require extra care in distribution design to avoid troublesome flicker. Motors larger than 6.5 hp are usually connected 3-phase. Three-phase service is not usually supplied in residential areas.

Light and Power from One Secondary Main. In a *radial* system, 3-phase service is sometimes supplied from a separate secondary main if voltage is affected by elevator motors or other intermittently used load. If separation of light and power service is not necessary, the nature of the connection may depend upon the relative size of light and power loads. When power load is predominant, lighting load may be served by providing additional capacity in one of the transformers and bringing in a neutral from it for the lighting service. The neutral for lighting service is sometimes derived from a transformer connected to one phase of 240- or 480-V power circuits giving 120/240 V for lighting. This is the usual procedure where power is served at 480 V. When the lighting load is predominant, service is often provided at 208Y/120V, 4-wire.

Transformer and Secondary-Main Economy, Overhead Distribution. Several independent studies have been made to determine the proper combination of transformer and radial secondary main that provides satisfactory voltage regulation and costs a minimum per kVA of load served. All these studies indicate that for 120/240-V single-phase distribution, overhead secondary mains should be three No. 1/0 to three No. 4/0 aluminum, the latter being preferred when air conditioning or heating is to be served.

Permissible length of the three No. 1/0 aluminum secondary mains depends on the load density. On the assumption of evenly distributed loads and 3% drop in the mains, for 15 kW/1000 ft, the permissible length is 600 ft, and for 30 kW/1000 ft, 400 ft. Widespread use of ranges and motor-driven appliances establishes an additional limit for flicker at 200 to 300 ft.

Transformer size should be such that the initial peak load is between 75% and 100% of rated capacity. In medium-load densities, 25- and 50-kVA transformers will fulfill this requirement. Transformers should be allowed to remain in service until their winter peak load reaches at least 150% to 180% of rated capacity. When this occurs, the “hot spot” winding temperature is approaching 110°C—the maximum safe temperature.

Load growth should be taken care of by *installing additional transformers* and cutting radial secondaries or by *increasing the size* of the existing transformers where secondary-main regulation permits. The three No. 1/0 to 4/0 aluminum single-phase secondary mains should not be replaced by larger conductors to improve secondary-main regulation. Additional transformers should be installed and parts of the existing mains transferred to the new transformers.

Underground systems should also be designed initially with capacity for growth. In order to accomplish this, many utilities in underground residential distribution (URD) work do not use secondary mains. Rather, one transformer is used to supply four to six homes by installing service drops large enough for future loads from the transformer to each home. With this system design it is relatively easy to change out the transformer to a larger size when the load grows.

Pad-mounted transformers can be sized and operated the same as overhead-type transformers. Advantage can be taken of the short-time overload capability given by ANSI C57.91, “Guide for Loading Mineral Oil Immersed Overhead-Type Distribution Transformers with 55C or 65C Average Winding Rise.” Subsurface transformers in close-fitting cylindrical vaults require special baffles and chimney specified by the manufacturer in order that they might be loaded the same as an overhead-type transformer.

Subway-type transformers should not be replaced or relieved of load until the calculated hot-spot winding temperature exceeds 110°C, provided, of course, that voltage at the ends of the secondary is satisfactory. To calculate hot-spot winding temperature, the maximum load and top-oil (or case) temperatures must be measured. Maximum case temperature has been found to be within 3°C of top-oil temperature. It is assumed in making the calculation that the difference between hot-spot-winding and top-oil temperature is 20°C at full load and that this difference varies as the square of load. This is a conservative assumption. For example, assume maximum case temperature 67°C when 130% load is on the transformer. Then the calculated winding hot-spot temperature is given by

$$67°\text{C} + 3°\text{C} + 20°\text{C}(1.30)^2 = 114°\text{C}$$

Fans to supplement natural air movement have been used to boost safe capability of vault transformers.

Table 18-10 gives the voltage drop per 10,000 A · ft for single-phase and 3-phase secondaries for a variety of load power factors. The underground portion of the table can be used for underground systems and also overhead systems where triplex cable construction is employed. The overhead part of the table gives the voltage-drop information for overhead aluminum conductors on racks. The table can be used to determine voltage drop quickly on any secondary circuit if load, circuit length, and conductor size are known.

All values in the table are for aluminum conductors at 50°C temperature. Values for copper can be determined with satisfactory accuracy by using the table for a conductor of equivalent resistance; that is, use an aluminum conductor two sizes larger than the copper conductor.

In using the table, the first thing required is the number of ampere-feet involved in the problem. This is obtained by multiplying the amperes per phase by length of circuit in feet. (For single-phase, use number of feet between source and load; impedance of return circuit is included in table.) Divide this ampere-feet by 10,000 to determine the multiplier to be used with values in the table. For the proper voltage, conductor size, and power factor, find the voltage-drop factor in the table and multiply by the multiplier determined previously. This will be the absolute line-to-neutral volts difference (drop) between the sending and receiving ends of the circuit. Dividing by the line-to-neutral voltage

TABLE 18-10 Voltage Drops per 10,000 A · ft* for Single-Phase and 3-Phase Secondaries, 60 Hz

	Voltage									
	120/240-V single-phase					208Y/120 V, 240 V, 480Y/277 V, and 480-V 3-phase				
	Lagging power factor									
Conductor size	0.7	0.8	0.9	0.95	1.00	0.7	0.8	0.9	0.95	1.00
Underground or triplex secondary‡										
Aluminum:										
No. 2	4.524	5.042	5.530	5.752	5.858	2.262	2.521	2.765	2.876	2.929
No. 1	3.690	4.084	4.450	4.606	4.646	1.845	2.042	2.225	2.303	2.323
No. 1/0	3.002	3.304	3.574	3.686	3.684	1.501	1.652	1.787	1.843	1.842
No. 2/0	2.458	2.684	2.880	2.954	2.920	1.229	1.342	1.440	1.477	1.460
No. 3/0	2.028	2.194	2.334	2.380	2.318	1.014	1.097	1.167	1.190	1.159
No. 4/0	1.684	1.804	1.898	1.920	1.840	0.842	0.902	0.949	0.960	0.920
350 kcmil	1.166	1.218	1.238	1.228	1.114	0.583	0.609	0.619	0.614	0.557
Overhead secondary‡										
Aluminum:										
No. 2	5.530	5.888	6.146	6.192	5.860	2.801	2.974	3.095	3.112	2.930
No. 1/0	3.932	4.088	4.150	4.102	3.700	2.002	2.074	2.097	2.067	1.850
No. 2/0	3.372	3.456	3.448	3.368	2.940	1.722	1.758	1.746	1.700	1.470
No. 4/0	2.516	2.504	2.406	2.284	1.840	1.294	1.282	1.225	1.158	0.920
336.4 kcmil	1.940	1.876	1.792	1.594	1.160	1.006	0.968	0.918	0.813	0.580
477 kcmil	1.646	1.556	1.392	1.248	0.820	0.858	0.808	0.718	0.640	0.410
795 kcmil	1.336	1.224	1.042	0.892	0.480	0.704	0.642	0.543	0.462	0.240

Note: 1 in = 25.4 mm; 1 ft = 0.3048 m. Regulation of copper conductors can be estimated with reasonable accuracy the same as that of aluminum conductors two sizes larger.

*Values in the table give the difference in absolute value between sending-end and receiving-end line-to-neutral voltages of balanced 3-phase circuit and phase-to-phase or phase-to-neutral voltages of single-phase circuit.

†Underground cable impedances are based on 50°C conductor temperature with close triangular spacing of cable using typical solid-dielectric insulation, 100% insulation level, single conductor.

‡Overhead conductor impedances are based on 50°C conductor temperature with 8-in equivalent spacing for single-phase and 10-in spacing for 3-phase.

of sending end or receiving end and multiplying by 100 will express this as a percentage of sending- or receiving-end voltage, respectively.

Example. Given a 3-phase 60-Hz secondary 500 ft in length, which consists of No. 4/0 aluminum conductor cable; conductor temperature 50°C; receiving-end load 100 kVA at 0.8 power factor lagging; receiving-end line-to-line voltage 480.

$$A \cdot ft = \frac{100}{\sqrt{3} \times 0.48} \times 500 \text{ ft} = 60{,}142, \text{ or } 6.014 \text{ times tabular value}$$

From Table 18-10 for No. 4/0 cable, 0.8 of the value is 0.902. Line-to-neutral voltage drop is 0.902 × 6.014 = 5.425. This is 5.425/277 × 100 = 1.96% voltage drop on basis of receiving end.

18.13 BANKING OF DISTRIBUTION TRANSFORMERS

Banking. Tying together the secondary mains of adjacent transformers supplied by the same primary feeder is known as *banking*. The practice of banking, when used, is usually applied to the secondaries of single-phase transformers, and all transformers in a bank must be supplied from the

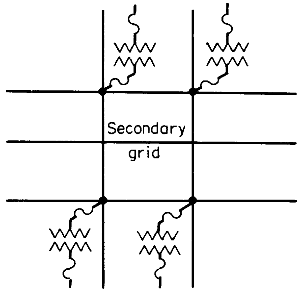


FIGURE 18-35 Fuse application in grid systems.

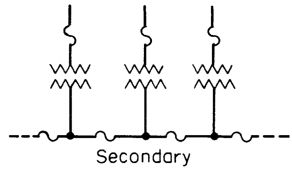


FIGURE 18-36 Fuse application in straight-line systems.

same phase of the primary circuit. The use of banking is not as prevalent as it was formerly. Banked distribution transformers differ from the low-voltage ac network in that one circuit supplies all transformers where secondaries are banked together, whereas different circuits supply adjacent transformers in an ac low-voltage network. Only a few companies operate their transformers banked.

Advantages claimed for banking compared with secondary radial distribution are (1) reduction in lamp flicker caused by starting motors; (2) less transformer capacity required because of greater load diversity among a larger group of customers; (3) better average voltage along the secondary; and (4) greater flexibility for load growth. There are two general types of secondary banking: the grid type and the straight-line type, as shown in Figs. 18-35 and 18-36.

18.14 APPLICATION OF CAPACITORS

Power Factor Correction. It is desirable to add shunt capacitors in the load area to supply the lagging component of current. The cost is frequently justified by the value of circuit and substation capacity released and/or reduction in losses. Installed cost of shunt capacitors is usually least on primary distribution systems and in distribution substations.

The application of a shunt capacitor to a distribution feeder produces a uniform voltage boost per unit of length of line, out to its point of application. Therefore, it should be located as far out on the distribution system as practical, close to the loads requiring the kilovars. There are some cases, particularly in underground distribution, where secondary capacitors are economically justified despite their higher cost per kilovar.

Development of low-cost switching equipment for capacitors has made it possible to correct the power factor to a high value during peak-load conditions without overcorrection during light-load periods. This makes it possible for switched capacitors to be used for supplementary voltage control. Time clocks, temperature, voltage, current, and kilovar controls are common actuators for capacitor switching.

Capacitor Installations. Capacitors for primary systems are available in 50- to 300-kvar single-phase units suitable for pole mounting in banks of 3 to 12 units. Capacitors should be connected to the system through fuses so that a capacitor failure will not jeopardize system reliability or result in violent case rupture. To ensure that the proper fuse protection is provided, the installed capacitor fuse ratings are listed in Tables 18-11 and 18-12 and the probability of rupture is shown in Table 18-13.

TABLE 18-11 Recommended Group Fusing, K- or T-Rated Links (Floating-Y Banks)

	3-Phase kilovar										
Volts	150	300	450	600	900	1,200	1,350	1,800	2,400	2,700	3,600
2,400	40K	—	—	—	—	—	—	—	—	—	—
4,160	25	40	65[a,b]	80K[a,d]	—	—	—	—	—	—	—
4,800	20	40	50[c]	—	—	—	—	—	—	—	—
7,200	12	25	40	50K[c]	80[a,e]	—	—	—	—	—	—
8,320	12	25	30	40	65[f]	80K[h]	—	—	—	—	—
12,470	8	15	25	30	50[g]	65[i]	65[a]	80K[j]	—	—	—
13,200	8	15	20	25	40	50	65[a]	80K[j]	100K[a,k]	—	—
13,800	6	12	20	25	40	50	65[a]	80[k]	100K[a,k]	—	—
14,400	6	12	20	25	40	50K	65[a]	80[k]	—	—	—
20,800	—	8	12	20	25	40	40	50	65	80[a]	100K[a]
21,600	—	8	12	15	25	30	40	50	65	80[a]	—
23,000	—	8	12	15	25	30	40	50	65	80T[a]	—
23,900	—	8	12	15	25	30	30	50	65	65	80K
24,900	—	8	12	15	25	30	30	50	65[j]	65	80K
34,500	—	—	8	10	15	20	25	30	40	50	65

Notes: Fusing is in safe zone unless otherwise shown. Max bank size for 50 kvar units is 600 kvar. Max bank size for 100 kvar units is 1200 kvar. Max bank size for 150 kvar units is 1800 kvar. Max bank size for 200 kvar units is 2400 kvar.

[a]Zone 1.
[b]150-kvar units only.
[c]Zone 1 for 50-kvar units.
[d]200-kvar units only.
[e]300-kvar units only.
[f]Zone 1 for 100- or 150-kvar units.
[g]Zone 1 for 100-kvar units.
[h]Zone 1 with 200-kvar units. Not suitable for 100-kvar units.
[i]Zone 1 for 100- and 200-kvar units.
[j]For 200-kvar and larger only, zone 1 for 200 kvar units.
[k]For 300-kvar and larger only.

Effect of Shunt Capacitors on Voltage. Proposed permanently connected capacitor applications should be checked to make sure that the voltage to some customers will not rise too high during light-load periods. Switched capacitor applications should be checked to determine that switching the capacitor bank on or off will not cause objectionable flicker. The curves in Fig. 18-37 can be used to compute voltage rise.

Effect of Shunt Capacitors on Losses. The maximum loss reduction on a feeder with distributed load is obtained by locating capacitor banks on the feeder where the capacitor kilovars is equal to twice the load kilovars beyond the point of installation. This principle holds whether one or more than one capacitor bank is applied to a feeder.

Capacitor kilovars up to 70% of the total kilovar load on the feeder can be applied as one bank with little sacrifice in the maximum feeder-loss reduction possible with several capacitor banks. A rule of thumb for locating a single capacitor bank on a feeder with uniformly distributed loads is that the maximum loss reduction can be obtained when the capacitor kilovars of the bank is equal to two-thirds of the kilovar load on the feeder. This bank should be located two-thirds of the distance out on the distributed feeder portion. Deviation of the capacitor bank location from the point of maximum loss reduction by as much as 10% of the total feeder length does not appreciably affect the loss benefit. Therefore, in practice, in order to make the most out of the capacitor's loss reduction and voltage benefits, it is best to apply the capacitor bank just beyond the optimum loss-reduction location.

TABLE 18-12 Recommended Group Fusing, K- or T-Rated Links

Grounded-Y- and Δ- Connected Banks

	3-Phase kilovar										
Volts	150	300	450	600	900	1,200	1,350	1,800	2,400	2,700	3,600
2,400	40	80	—	—	—	—	—	—	—	—	—
4,160	25	50	80	100	—	—	—	—	—	—	—
4,800	20	40	65	80	140	—	—	—	—	—	—
7,200	15	30	40	65	80	—	—	—	—	—	—
8,320	12	25	40	50	80	100	—	—	—	—	—
12,470	8	15	25	40	50	65	80	100	140	—	—
13,200	8	15	25	30	50	65	80	100	140	—	—
13,800	8	15	25	30	50	65	65	100	140	140	—
14,400	8	15	20	30	40	65	65	80	140	140	—
20,800	6	10	15	20	30	40	50	65	80	100	140
21,600	6	10	15	20	30	40	40	65	80	80	140
23,000	6	10	15	20	25	40	40	50	80	80	100
23,900	6	8	12	20	25	40	40	50	80	80	100
24,900	6	8	12	15	25	40	40	50	65	80	100
34,500	6	6	10	12	20	25	25	40	50	50	80

Notes:
1. Refer to Table 18-13 for fuse sizes within fault current limits.
2. Maximum link size for each unit—check Table 18-9 for all:

50 kvar	65K, 30T	Check Table 18-3
100 kvar	80K, 50T	Check Table 18-3
150 kvar	100K, 50T	Check Table 18-3
200 kvar	100K, 65T	Check Table 18-3
300 kvar and up	140K, 80T	Check Table 18-3

3. Ratio of fuse continuous current rating to nominal capacitor current is 1.65 minimum.

TABLE 18-13 Coordination Table: Grounded-Y and Δ Connected Banks

Maximum fault current for zone indicated.

	50 kvar unit		100 kvar unit		150 kvar unit		200 kvar unit		300 and 400 kvar unit	
Fuse link	Safe zone	Zone 1	Safe zone	Zone 1	Safe zone	Zone 1	Safe zone	Zone 1	Safe zone	Zone 1
30 K and lower	2900	3900	4000	5300	4600	6300	5400	7000	5800	7000
40 K	2700	3900	4000	5300	4600	6300	5400	7000	5800	7000
50 K	2000	3700	3900	5300	4600	6300	5400	7000	5800	7000
65 K	—	2400	2800	5300	4000	6300	5400	7000	5800	7000
80 K	—	—	700	3500	2200	5500	4100	7000	5000	7000
100 K	—	—	—	—	—	2800	1700	6300	2800	7000
140 K	—	—	—	—	—	—	—	1800	—	3500
20 T and lower	2900	3900	4000	5300	4600	6300	5400	7000	5800	7000
25 T	2200	3900	4000	5300	4600	6300	5400	7000	5800	7000
30 T	800	2800	3200	5300	4200	6300	5400	7000	5800	7000
40 T	220	1000	1700	4300	3000	6300	4500	7000	5600	7000
50 T	—	200	400	2500	1100	4000	2800	7000	4200	7000
65 T	—	—	—	500	—	2100	1600	5500	2500	6800
80 T	—	—	—	—	—	—	—	3500	1000	5000
100 T	—	—	—	—	—	—	—	—	—	2200

Note: Safe zone—rupture probability less than 10%. Zone 1—rupture probability 10% to 50%.

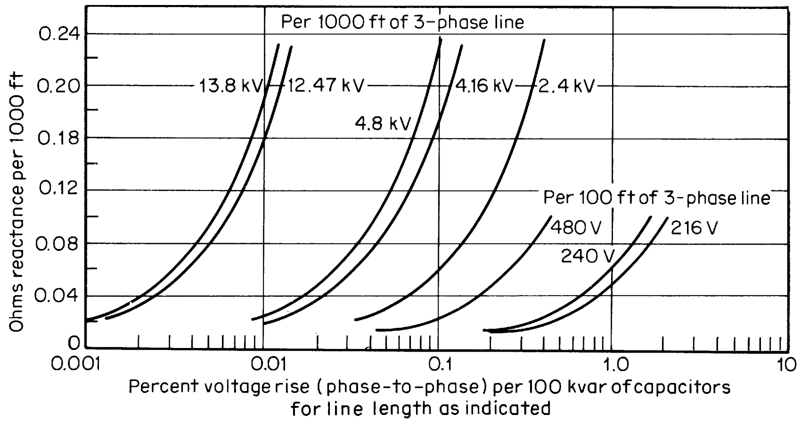


FIGURE 18-37 Curves of voltage rise caused by capacitor application.

18.15 POLES AND STRUCTURES*

Overhead Construction. Although overhead distribution construction remains less costly than underground, the vast majority of *new* residential developments are being served by underground systems. However, most new main feeder circuits and rural and semirural systems are being constructed overhead. Also, because of the tremendous amount of overhead distribution plant already in place, it is important to have a good grasp of overhead construction.

Appearance Considerations. For many years, the traditional overhead distribution construction embodied one or more crossarms on each pole for mounting primary insulators, distribution transformers, surge arresters, cutouts, etc. However, back in the early 1960s when public interest in appearance made itself manifest, many new concepts in overhead construction were introduced in the interest of obtaining better-appearing systems. The principal change was the introduction and wide use of "armless" construction wherein insulators and equipment are directly mounted to the poles. Other ideas also were introduced, such as use of shorter poles, cabled secondaries and services, strategic use of steel poles to reduce or eliminate need for guy wires, and fewer circuits per pole. It should be noted, however, that cabled secondaries employ a "covered conductor" which can withstand momentary contact between conductors but which is appropriately considered by the National Electrical Safety Code ANSI C2 to be the same as bare conductor for clearances to other objects and for all other purposes. In addition, the black covering adds both thickness and contrast to the cables when viewed against the sky.

Wood poles have been used almost universally for overhead distribution lines because of the abundance of the material, ease of handling, and cost. The life of wood poles is materially extended with wood preservatives, and cedar, pine, and fir are most commonly used.

Specifications and dimensions for wood poles are presented in ANSI 05.1. Poles meeting the requirements of this standard are grouped into classes based on their circumference at a location 6 ft from the butt. Poles of a given class and length are designed to have approximately the same load-carrying capacity regardless of species. ANSI 05.1 used a uniform methodology for testing poles that

*To avoid inappropriate duplication, certain discussions common to transmission and distribution may be found in Sec. 14.

required a uniform fulcrum-point location. This point generally was 2 ft plus 10% of the length of the pole measured up from the butt of the pole. This same value has been used as a setting depth for many locations but is not sufficiently deep for some soils.

The minimum circumferences specified at 6 ft from the butt have been calculated in order for each species in a given class to develop, at the ground line, appropriate stresses when a given horizontal load is applied 2 ft from the top of the pole. The horizontal loads used in the calculations for identifying the 15 classes are as given in Table 18-14.

TABLE 18-14 Loads Used to Identify Pole Classes

Pole class	Horizontal load, lb
H6	11,400
H5	10,000
H4	8,700
H3	7,500
H2	6,400
H1	5,400
1	4,500
2	3,700
3	3,000
4	2,400
5	1,900
6	1,500
7	1,200
8	740
10	370

Note: 1 lb = 0.4536 kg.

In making the calculations, it was assumed that the pole is used as a simple cantilever and that the maximum fiber stress in the pole due to the bending moment will occur at the ground level. The circumference at the ground line was calculated using standard engineering formulas. This circumference value was then used to calculate the circumference 6 ft from the butt using typical tapers per foot of length between the ground line and the point 6 ft from the butt.

The assumption of maximum fiber stress at the ground line is theoretically correct if the taper of the pole is such that the circumference at the ground line is not more than 1$^1/_2$ times the circumference at the point of loading. If the circumference at ground line is more than 1$^1/_2$ times the circumference at the point of loading, the maximum fiber stress theoretically occurs at a location above the ground line where the circumference is 1$^1/_2$ times the circumference at the point of loading. This makes it necessary to calculate specific loads supported.

Typical tapers used in determining the required circumference 6 ft from the butt circumference range of 0.38 for western cedar to 0.20 for western hemlock. These numerical values are in change in circumference per foot of length.

Concrete poles reinforced with steel originally were employed chiefly for street-lighting standards, where a neat appearance is demanded. But some concrete poles have been used for general distribution as well, usually with a minimum of attached wires and apparatus.

With the increased manufacturing and quality control capability for prestressed concrete poles, and the need for tall structures for narrow rights of way or aesthetic requirements, has come the more frequent use of prestressed concrete poles in transmission lines. As demands for transmission facilities in urban narrow rights of way have increased, they have been most often met with steel single-pole structures, especially where two circuits are required.

Steel poles, ordinarily set in concrete, have long been used to support street lights. More recently, in a more ornamental form and bolted to concrete foundations, they have been used for parkway lighting and, to a limited extent, for distribution where appearance demands.

Aluminum poles also are employed for parkway lighting standards and for certain other locations. They are bolted to concrete foundations to avoid the attack of fresh cement on aluminum. Although use of aluminum on transmission lines has increased, it has been with latticed structures, not generally with poles.

Types of Loading. Poles carrying overhead distribution lines are subject to vertical and horizontal forces, of which some are continuous and others are applied only under abnormal or occasional conditions. Normal vertical forces are the weight of wires, transformers, and other equipment, and these are less than normal horizontal forces in many cases. Abnormal vertical load is imposed when wires are coated with ice, which may increase their normal weight 200% to 400%. For example, the weight of six covered No. 6 copper wires 100 ft long is normally 67 lb, but ice to a radial thickness of 0.5 in increases their weight to about 370 lb.

Normal horizontal forces acting on a pole are the unbalanced component of wire tension at turns and corners, the side pull of service drops, and the horizontal component of weight when the pole is not vertical. Abnormal horizontal stresses are imposed by wind pressure, by breakage of conductors, or by failure of supporting guys.

Application of Loading. Vertical loading of wires and equipment is applied through crossarms and other attachments to the pole. These forces are amply sustained by poles chosen to meet requirements of transverse forces, except that for line transformers, poles having 1-in greater diameter than line poles may be chosen. Transverse forces from unbalanced conductor tension at corners and bends are normally the greatest forces acting on the line. These are usually carried by guy cables secured to suitable anchorages, which relieve the pole itself of the stress. In some cases, the pole is underbraced and carries the entire load.

Ice Loading. When wires are loaded with ice, conductor tension is increased in direct proportion to the added weight of ice and may become two to four times as great as normal. This stress is borne by the conductors and, through them, communicated to the pole and the guying system. Where ice loading occurs, the guying system must have a suitable factor of reserve to meet abnormal loadings. The tension of conductors being increased with ice loading, elasticity in the wire permits a slight increase in length which makes tension less than the calculated amount for nonelastic conductors and supports.

Wind Loading. Loading due to wind pressure becomes appreciable in the design of poles and structures when wind velocities of over 40 mi/h are prevalent. Such forces are most noticeable on overhead lines when the direction of wind is at right angles to the direction of wires, both because the area exposed is greatest at that angle and because the force exerted is sustained by the pole without the aid of guying.

The area of conductors exposed to wind is much increased by a coating of ice, and the combination of ice with high wind is often the most severe loading condition to which a line is subjected. In many parts of the United States such a condition is never experienced, and it is very rare even where ice coatings occur almost every winter, since the conductor movement due to wind tends to break off the ice coating.

However, with the introduction of large-diameter conductors and bundled-conductor configurations, high winds in summer or other storm periods not involving ice may cause the greater problem. As a result, the NESC requires that severe wind conditions be checked if any portion of the structure or conductors exceeds 60 ft above grade.

Strength of Wood Poles. The strength of a wood pole must be sufficient to withstand transverse forces such as wind pressure on the pole and conductors, unbalanced pull on conductors when they are broken, and side pull on curves and corners where guys cannot be provided. These forces place the fiber on the wood under tension, and the load which a pole will carry is determined by the inherent strength of its wood fiber under tension and the moment of forces. The moment is

$$M = PL + P_1L_1 + P_2L_2 + \cdots \tag{18-13}$$

in which P = force, lb, acting at one crossarm, L = height, ft, at which the arm is attached, P_1, P_2, etc. = forces acting on other arms, and L_1, L_2, etc. = respective heights. If s = fiber stress, lb/in^2, and c = circumference at the ground, in, the allowable moment of a pole of given size is

$$M = 0.00026386sc^3 \tag{18-14}$$

Thus, for a pole having a ground-line circumference of 40 in and an allowable fiber stress in an emergency of 2500 lb/in^3, the maximum allowable moment is

$$M = 0.00026386 \times 2500 \times (40)^3 = 42{,}218 \text{ lb} \cdot \text{ft}$$

If the average height of attachments is 30 ft, total force is 42,200/30 = 1407 lb.

The ultimate fiber stresses for various species of wood poles are listed in ANSI 05.1. In practice, the actual pole stresses are limited to some allowable percentages of ultimate stress. The NESC, provides guidelines for allowable stresses under vertical and transverse loading. The National Weather Service now records wind data that include gust speeds. Consult the latest *NESC Handbook* for requirements in effect at the time of line design. Some utilities use larger factors of safety in their designs to allow for errors in tensions, nonuniformity of soils, or special loading requirements; different factors of safety may be used for normal unbalanced forces and for abnormal forces of a temporary nature.

When the maximum fiber stress is above the ground line because of taper of the pole, the allowable moment can be calculated in a manner similar to Eq. (18-14) using the circumference of the pole at the point of load application and the taper of the pole to determine the location where the circumference is $1^1/_2$ times that at the location at the load. Usually decay is greatest at the ground line, and as the pole ages, its ground-line circumference is reduced to make it the point of greatest fiber stress.

Example of Calculation of Pole Size. ANSI 05.1 specifies a value of ultimate fiber stress for western cedar of 6000 lb/in^2. Assume a 40-ft pole and a depth of setting of 6 ft. Assume the pole will carry a transverse pull of 400 lb at a height of 32 ft and another of 90 lb at 30 ft. What class of pole is required and what are its circumferences at ground line and at top?

$$M = PL + P_1L_1 = 400 \times 32 + 90 \times 30 = 15{,}500 \text{ ft} \cdot \text{lb}$$

If a factor of safety of 4:1 is used with respect to the ultimate stress, the allowable fiber stress is 6000/4 = 1500 lb/in^2. From Eq. (18-14),

$$c = \sqrt[3]{\frac{M}{0.00026386s}} = \sqrt[3]{\frac{15{,}500}{0.00026386 \times 1500}}$$

$$= \sqrt[3]{39{,}162} = 33.96 \text{ in at the ground line}$$

From ANSI 05.1, Table 5, this corresponds very closely to a Class 5 pole, which has a circumference of 34.0 in (10.82 in diameter) at 6 ft from the butt. Minimum circumference of this class of pole at the top is 19 in (6.05 in diameter).

If the allowable fiber stress were increased to one-half the ultimate (factory of safety = 2), the moment could be increased to

$$M = 0.00026386 \times 3000 \times (34)^3 = 31{,}112 \text{ ft} \cdot \text{lb}$$

If this were a northern white cedar pole, having an ultimate fiber stress of 4000 lb/in^2 and a safety factor of 4, the ground-line circumference would be

$$c^3 = \frac{15{,}500}{0.00026386 \times 1000} = 58{,}743 \text{ ft} \cdot \text{lb}$$

$$c = \sqrt[3]{58{,}743} = 38.87 \text{ in}$$

From Table 3 of ANSI 05.1, this is very close to the Class 5 pole, which has a specified circumference of 39 in 6 ft from the butt. This also illustrates that the pole classification, in effect, defines the loading capability of the pole regardless of the species.

Wind Pressure. Wind pressure must be taken into account when designing a pole line. For purposes of calculation, the following formulas are often used to calculate pressures due to wind:

$$P = 0.004V^2 \tag{18-15}$$

where P = pressure, lb/in^2, on flat surfaces normal to the wind, and V = wind velocity, mi/h

$$P = 0.0025V^2 \tag{18-16}$$

for cylindrical surfaces such as wires and poles. Values of V can be obtained from weather bureau records for the particular locality.

Wind pressure on a 40-ft pole, of which 34 ft is above ground, with 7 in top diameter and 14 in butt diameter, can be calculated as follows:

$$\text{Projected area} = [(7 + 14) \times \tfrac{1}{2}] \times 34 \times 12 = 4284 \text{ in}^2$$

$$= 4284/144 = 29.75 \text{ ft}^2$$

A wind of 60 mi/h would cause a force calculated by Eq. (18-16):

$$P = 0.0025 \times (60)^2 = 9.0 \text{ lb/ft}^2$$

Since the uniform wind pressure is applied to a long, slender trapezoidal area whose center of gravity is 15.11 ft above the ground line, the resulting moment about the ground line is

$$9 \times 29.75 \times 15.11 = 4046 \text{ ft} \cdot \text{lb}$$

With the diameter of a typical distribution conductor taken as about 0.35 in, the total wind force on a 150-ft span assuming 60 mi/h would be

$$\left(150 \times \frac{0.35}{12}\right) \times 0.0025 \times (60)^2 = 39.375 \text{ lb}$$

On six conductors it would be 236.25 lb. Assuming that the conductors have an effective height of 31 ft at the pole, the resulting moment is $236.25 \times 31 = 7324$ ft · lb. The sum of the moments from wind pressure on pole and wires is

$$4046 + 7324 = 11{,}370 \text{ ft} \cdot \text{lb}$$

When distribution transformers, capacitors, voltage regulators, or other equipment are mounted on the pole, resulting wind forces should be taken into consideration.

In some areas where (1) overhead facilities are subject to high winds; (2) the soil is such as to provide relatively poor overturning resistance; and (3) underground facilities are inappropriate, such as barrier islands subject to hurricane and tidal forces, a unique system is employed to provide for minimum storm damage and quick return to service. In these cases, the structures are deliberately overdesigned relative to their natural foundations so that they will lean, rather than break, under excessive wind loading. Power is turned off as winds reach hurricane speed. After the storm, the lines are inspected, leaning poles are straightened, and service is restored, all with a minimum outage time and expense.

Ice- and wind-loading requirements vary widely throughout the United States, and each utility has adopted design practices suitable for its own conditions of terrain and climate. The NESC, presents some guidelines on conductor loading.

The actual loading on conductors is equal to the resulting loading per foot due to the vertical load on the conductor, assumed ice-covered where appropriate, and the transverse loading due to horizontal wind pressure on the projected area of the conductor, again assumed ice-covered where appropriate. Usually a design constant is added to the loading so calculated.

To establish the guidelines for the loading of overhead lines, the NESC has divided the United States into three loading districts, heavy, medium, and light. These are roughly (1) the northeastern section extending east to west from the Atlantic through the Dakotas and from the Canadian border southward into Texas and the Ohio Valley; (2) the Pacific Northwest and a narrow belt eastward to the Atlantic including parts of Arizona, Texas, Louisiana, Alabama, Georgia, and the Carolinas; and (3) the remaining narrow belt extending from coast to coast and bordered on the south by Mexico and the Gulf of Mexico. These are, respectively, the heavy, medium, and light loading districts.

TABLE 18-15 Conductor Loadings Due to Ice and Wind

	Loading district			
	Heavy	Medium	Light	Extreme wind loading
Radial thickness of ice, in	0.50	0.25	0.00	0.00
Horizontal wind pressure, lb/ft^2	4	4	9	9 to 31
Temperature, °F	0	+15	+30	+60
Constant to be added to the resultant, lb/ft	0.30	0.20	0.05	0.00

Note: 1 lb/ft = 1.488 kg/m; 1 lb/ft^2 = 4.882 kg/m^2; $t_{°C} = (t_{°F} - 32)/1.8$. Selected data from the National Electrical Safety Code, IEEE C2-1997. Since heavy ice does not often form on conductors in a heavy wind, the transverse loading assumed is deemed sufficient for the purpose but is not sufficient to represent the vertical (or combined) load which is imposed on conductors by heavy deposits of ice which frequently form in comparatively still air. In order to apply a total loading to conductors representing more nearly the conditions encountered in practice, constants are added to the conductor loading.

Exact locations of these districts may be found in the *NESC Handbook*. Table 18-15 illustrates suggested loadings. The constant is used only for conductor strength checks. It is not used in applying loads to structures.

Equipment Loading. Poles are subject to heavy loads when used to support equipment such as distribution transformers, voltage regulators, and capacitor banks. Such loads are chiefly vertical but do have a transverse component when the pole is bent or drawn away from a vertical position. The equipment also presents additional transverse loading due to wind pressure. Therefore, poles to be used for supporting equipment usually are specified to be of a better class than those supporting conductors only.

Three-unit installations of distribution transformers normally employ the "cluster-mounting" arrangement in which the transformers are supported directly on the pole by suitable brackets. Three-phase banks as large as 500 and 750 kVA are installed in this manner. The former practice of supporting the transformers on a platform carried by two or more poles has nearly disappeared because of appearance and cost considerations.

Unbalanced Loads. Transverse forces are imposed on a pole where there is a change of direction of the line, that is, where the conductors on either side of the pole form an angle. These forces usually are offset by guys where practicable. The loading on the structure, including guys, is considered to be the resulting load equal to the transverse wind load and the load imposed by the conductors due to the change in direction. If it is not practical to guy, the pole and its setting must have sufficient strength to withstand the stresses imposed. The force applied to the pole at the point of attachment of the conductors can be calculated as in Fig. 18-38. Divergence from a straight line can be determined by joining two points, each 100 ft from the corner pole, by a straight line. The distance A from the corner pole perpendicular to this line can then be used to calculate the transverse force applied by the line wires on the pole.

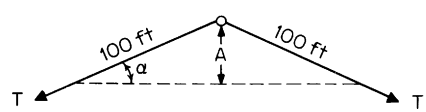


FIGURE 18-38 Determining transverse force on a corner pole.

Assume T to be the total tension, in pounds, of all the wires carried by the pole.

$$\text{Transverse force} = 2T \sin \alpha = 2T \frac{A}{100} \tag{18-17}$$

For example, if A is 10 ft, and if there are six wires having a total tension of

$$T = 6 \times 250 = 1500 \text{ lb}$$

the transverse force due to conductor tension to be sustained normally by the pole is

$$2 \times 1500 \times \frac{10}{100} = 300 \text{ lb}$$

For right-angle turns, α is 45°, $\sin \alpha = 0.707$, and $A = 70.7$ ft. This is the most severe condition ordinarily encountered. In such cases, tensions should be reduced as much as practicable at the corner pole by shortening spans and by transferring a part of the stress to guy wires. Stresses at dead-end poles are treated similarly.

Safety-Code Requirements. The National Electrical Safety Code, in Part 2, sets up minimum safety requirements for loading, strength, and clearance for those parts of supply lines which are involved in crossings of railroad or communication circuits or which come into such proximity to these as to create a "conflict." It also includes joint use of lines and separate use on any public way. The code recognizes differences in degree of estimated hazard, which are assumed to depend on voltage of the supply line, importance of railroad and other communication systems, classification of loading district, types of crossings, etc. The NESC allows reduced loadings where utility facilities are sheltered from the wind, but it requires full loadings to be met if those shelter mechanisms do not exist. Caution is advised in depending on the continued existence of buildings and other artificial objects as a sheltering influence—especially in the case of buildings; they can either (1) be removed or (2) be joined by others and produce channeling effects that would adversely affect the line loading. Similarly, trees are not generally considered to be an effective shelter because of the possibility of clear-cutting or other removal.

18.16 STRUCTURAL DESIGN OF POLE LINES

Pole Location. In residential areas, poles are generally spaced from about 100 to 150 ft apart depending on the size of the lots. The span usually is an integral number of lot widths in length, and poles are set to provide convenient points for connection of services. Longer spans up to 300 ft or longer are used in rural areas. In either case, pole spacing should consider expected future growth and service requirements in the area. The choice of pole spacing is also a function of the load to be carried and the relative economics of longer spans. Pole spacing and line location both are a function of the obstacles in the area. The NESC, ANSI C2, has specific requirements about placement of poles near other poles, roads, buildings, fire hydrants, etc.

Poles are required to be at least 3 ft away from fire hydrants and should be at least 4 ft away. Although these distances are required to allow firefighters access in emergencies, they also give room for personnel working on the pole to move around on the ground. No fire hydrant, telephone pedestal, or other like object should be located on the climbing side of the pole; the climbing side of the pole should be kept clear of protruding obstacles to provide a clear drop zone for line workers if they chip out of the pole and fall.

There are two hazards with the line-worker's gaff cutting out chipping on a pole; the first is the possibility of injury due to picking up splinters from previous gaff marks, and the second is the possibility of injury from the fall. Because of the severe problems caused from splinter injury, some communication line workers (who do not climb as high as supply workers) are taught to climb a pole without belting off until they reach the work height; this allows them to push away from the pole as they go down and allows them to roll when they hit the ground. In such cases, no obstruction should be allowed on the climbing side within 10 ft.

Poles and their supported equipment up to 15 ft above a roadway are required to be located far enough from roadways so that an ordinary vehicle that is using and located on the traveled way will not contact the utility facilities.

Poles are required to be at least 12 ft from railroad rails, except in some limited sidings where the clearance must be 7 ft, and room must be left to unload cars. Where some other facility is the controlling obstruction, the clearance may be reduced from 12 ft but may not be less than 7 ft.

Selection of Poles. The height of poles is determined by required clearances over obstructions, streets, and crossings, the span lengths, and the number and character of conductors or circuits to be carried. The most usual lengths for poles used in distribution construction have been 30, 35, and

40 ft. Armless construction and joint use tend to favor the 40-ft length. In recent years, however, as space for CATV and other facilities is provided on joint-use poles, 45-ft poles are even more commonly used, especially in hilly terrain. The Class 5 and 6 poles are very popular, although larger poles are often required where heavy equipment is to be mounted or longer spans with greater wind loadings are used. The class of pole is determined by stress requirements for the grade of construction being employed.

Where primary conductors are to be carried with joint use, the 35-ft pole usually is the minimum used, and 40 ft is quite common. Without primary conductors and with joint use, 30-ft poles are used occasionally. Joint use of pole lines by two or more public service companies is encouraged where feasible because (1) it makes maximum use of capital investment and avoids undesirable duplication of pole lines, and (2) it provides fewer opportunities for one circuit to fall into another during a storm. When a pole or section of poles is broken as a result of the action of falling trees, flying debris, or errant vehicles, the tendency is for all circuits of a joint-use line to be promptly and automatically deenergized. However, if a conflicting line falls into another line, facilities or personnel working on the other line may be damaged. Thus conflicting line locations are discouraged where joint-use locations are practical. In either case, the NESC, requires greater strengths than for single-circuit installations.

TABLE 18-16 Depth of Wood-Pole Settings

Length of pole, ft	Depth of setting, ft
30	5.5
35	6
40	6
45	6.5
50	7
55	7.5
60	8
70	9

Note: 1 ft = 0.3048 m.

Pole-Setting Depths. Table 18-16 lists typical depths of setting for wood poles of various lengths. These are sufficient for most soils but must be increased for larger poles set in lighter or more fluid soils; these setting depths are based on the fulcrum points used for testing pole strength in ANSI 05.1. They do not vary with pole diameter and may not be sufficient for large poles with large loads.

Guying Longitudinal, Angle, and Dead-End Force with Tensile Guys and Anchors. When the horizontal loads to be carried by poles are greater than can be safely supported by the poles, guys or braces are required to provide additional support. Guys and anchors are commonly used wherever conductor tensions are not balanced, as at dead ends, corners, or where the direction of the line changes substantially. Down guys transmit force from the overhead structure system to a buried anchor system (Fig. 18-39). They are located opposite the forces and use materials in tension to balance the forces. Where it is not practical to place these facilities to continue the imbalanced forces to ground, a compressive guy or pole brace must be used.

Where traffic ways or other obstacles do not allow direct anchoring of the forces at the imbalanced pole, an overhead span guy is used to transmit the force to another pole in line with those forces and in a place that allows a down guy to be used. The best down guy is a straight anchor guy composed of appropriate wire, fastenings, and anchor; it is the most economical to install in most locations, is the least trouble to maintain, deteriorates the least over time, and is the most reliable type of down guy.

The NESC, requires guys to have 4.72 m (15.5 ft) clearance over roadways and 2.89 m (9.5 ft) over spaces or ways accessible to pedestrians only. Where an anchor guy cannot be used because of interference with pedestrian or vehicular traffic, and where there is no practical location for an auxiliary pole and anchor guy that would allow a span guy to be used, a stub guy may have to be used. A stub guy consists of an angled span guy from the imbalanced pole to a short stub pole that is (1) high enough to provide the clearance required for the guy over the affected area and (2) low enough that the resulting moment arm is short enough that the stub can take the forces involved. Often the stub pole will have to be separately braced below ground to take the forces. Obviously, these conditions are not desired, and they are not practical for balancing large forces.

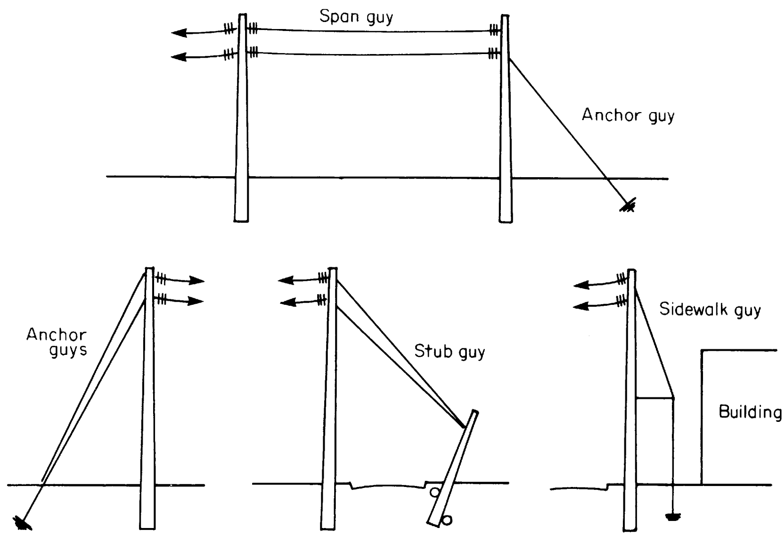


FIGURE 18-39 Tensile guys.

Another down guy used in constrained locations is the sidewalk guy. With the sidewalk guy, a horizontal strut is attached to the pole midway down the pole; the guy wire runs angled from the top of the pole to the strut and straight down to an anchor from there. These guys are most often used to guy small taps off a main line where there is not room for a full-length anchor-guy lead. The strut is usually just long enough to place the anchor back of a sidewalk and high enough to allow pedestrian clearance, hence the name. This guy should not be confused with the compressive guy discussed below; both are sometimes called *strut guys*. Sidewalk guys cannot be used for large forces because they will bow the pole due to their horizontal force; this, in turn, causes serious vertical-loading and bending-moment problems if the forces are great enough.

Where unbalanced tension is experienced by crossarms, such as in sidearm construction or dead ends for heavy conductors, crossarms are often guyed. Often a span guy will be attached to the back of the crossarm at each conductor location and run to another pole in line with the forces; otherwise, excess vertical loading on the crossarms can result.

Guy wires generally are made of stranded steel cable, usually galvanized for weather resistance. Several grades are available, including extra-high strength, which is commonly used. It is available in several diameters in steps of $^1/_{16}$ in from $^3/_{16}$ in up.

The National Electrical Safety Code, ANSI C2-2002, specifies that guy wires be used so that, with the required assumed loads and overload factors, they will not be stressed beyond 90% of ultimate strength, and for dead ends, not beyond 66.67% of ultimate strength. After the tension has been calculated, the guy cables are selected so that these requirements are met. The anchor guys usually are installed at a distance from the pole not less than one-quarter or more than 1 to $1^1/_2$ the height of the guy attachment. Generally a 1:1 slope is preferred. This lessens both the tension in the guy wire itself and the vertical component of forces transmitted to the pole. Where steep angles are used for the guy, or where large forces are otherwise involved, the pole size may need to be increased in order to withstand the vertical forces. In poor soils, this can even push the pole into the ground over time, thus decreasing clearances and twisting the pole, the latter because the now-excessive length of the guy allows the pole to move.

Compressive Guys and Pole-Foundation Underbracing. Where poles with imbalanced forces are so located that tensile guys cannot practically be used, either the forces must be taken by a compressive guy or the pole and its foundation must be made strong enough to take the forces by themselves. A *compressive guy* consists of a pole (or other member capable of taking the imbalanced load) placed at an angle leaning into the imbalanced pole and attached at the top so that it pushes against the force to be balanced. Its common name is a *push guy* for this reason; it is also called a *strut guy* in a few locations—not to be confused with the sidewalk guy. The most common use for a push guy is on the inside of a curve on a mountain road where, in essence, there is no land on the other side of the road in which to place a pole and anchor to use a span guy; span guys are preferred where they are practical because of their relatively simple installation, longer life, and ease of maintenance compared with the push guy. Push guys are difficult to attach to the supported poles, often can only be attached significantly lower than the imbalanced conductors, especially on sharp angles, and require the imbalanced pole to be oversized as a result. It is generally difficult to position the push guy in the ground, or cut it to the right length, so that the imbalanced pole will remain upright under load.

Where poles supporting unbalanced stresses are so situated that it is not practical to support them by either tensile or compressive guying, they must be underbraced to withstand the force imposed with as little deviation from original position as possible. This normally requires that the pole have more than usual diameter and be no taller than absolutely necessary for clearance.

For very long spans or heavy conductors, steel poles are often required. If wood poles are used, top diameters of 8 to 10 in are required for such positions to avoid bending. In addition, the pole is underbraced by timbers bolted to it below ground line and at the butt, as shown in Fig. 18-40. An alternative method, using concrete, is also good. Small boulders or crushed stone may be used in backfill to advantage.

In the use of plank or concrete, the pole is set at a slight angle, or rake, in a direction opposite that from which stress is to be applied, to allow for compacting of soil when wires are pulled up. The area of plank or concrete should be about 4 ft^2, both top and bottom. Where steel poles are used to secure strength, they are usually set in a concrete base of such dimensions as to bear the stresses imposed.

Vertical Clearances above Grade. Rule 232 of the National Electrical Safety Code, ANSI C2-2002, specifies the vertical clearances that utility wires, conductors, and cables must maintain above railroads, roads and other ground areas, and water areas. Before 1990, these clearances were specified when the conductor was at 15.5°C (60°F) conductor temperature (not ambient air temperature) and at specified ice loadings, whichever produced the greater sag. For span lengths greater than the basic span length applicable for the ice-loading area (heavy, medium, or light) and for conductor temperatures above 49°C (120°F), addititional clearances were required.

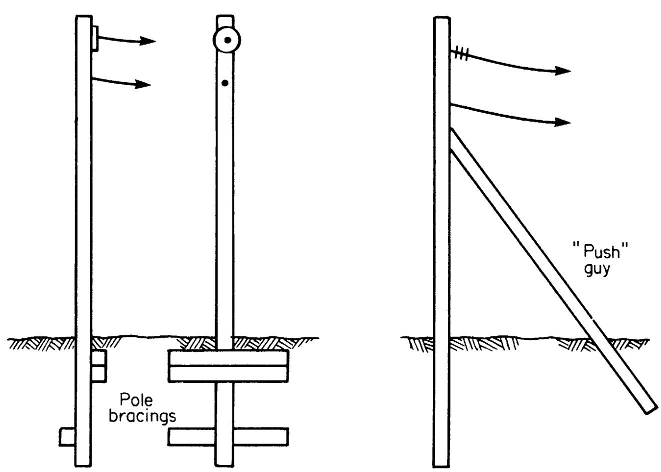


FIGURE 18-40 Comprehensive guy and pole bracing.

In 1990, the National Electrical Safety Code respecified the clearances to apply at maximum-sag conditions and coordinated the clearances with a system of reference dimensions of expected activity under the line and building blocks related to actual or potential voltage levels. For roads, parking lots, and other areas expected to have trucks, the reference dimension is 4.27 m (14 ft). For areas limited to personnel or restricted traffic only, the reference dimension varies from 2.4 to 3.0 m (8 to 10 ft), depending on the expected problem.

The building blocks include 30.42 cm (1.0 ft) for structure arms, 45.72 cm (1.5 ft) for neutrals and communication cables and grounded guys, 0.6 m (2.0 ft) for secondary bushings and jumpers up to 750 V and duplex, triplex, and quadruplex cables with bare neutrals, 0.76 m (2.5 ft) for open (bare or covered) supply conductors up to 750 V, 1.22 m (4.0 ft) for primary voltage bushings up to 22 kV, and 1.37 m (4.5 ft) for primary voltage conductors up to 22 kV. Additional clearances apply above 22 kV or 1 km (3300 ft) above sea level.

The previous 15.5°C (60°F)–based clearance requirements included 0.46 m (1.5 ft) to allow for sag change from 15.5 to 49°C (60 to 120°F) or specified ice-loading conditions. The 5.6-m (18.5-ft) minimum clearance specified in the NESC beginning in 1990 for distribution primary voltage conductors above a road 4.3-m (14-ft) reference dimension plus 1.4-m (4.5-ft) building block produces the same essential clearance requirement as the previous system. Similarly, the minimum clearance for a distribution neutral, guy, or communication cable is 4.7 m (15.5 ft, i.e., 14 + 1.5).

The present system (1) requires the installer to use actual expected changes in sag from the installed sag to the maximum final sag expected; (2) better states the real clearance involved; and (3) does not include the previous clearance penalties for short spans with less than 0.46 m (1.5 ft) of sag change from its 15.5°C (60°F) position.

It is important to stress that the NESC vertical clearances apply when the conductor is at its maximum sag position, even including emergency sag conditions, such as when one power line picks up the load from another line and the line losses heat the conductors above their normal maximum temperature. If special, localized icing conditions apply to the line, they also should be used to determine whether the summer or winter maximum sag conditions will be determinant.

Clearances Between Crossing or Adjacent Wires, Conductors, or Cables Carried on Different Supporting Structures. Where any two wires, conductors, or cables cross or are adjacent in open span without being supported on a common structure, the clearances between them must be sufficient to limit the possibility of contact between them. The effect of movement under wind, thermal, or ice loading must be considered as well as the required clearances between them. The National Electrical Safety Code considers that the two suspended facilities must first be considered to be in their most proximate position relative to each other, assuming that both experience the same ambient conditions (i.e., wind direction and speed, air temperature, and icing conditions); then the distance between them must meet the clearance requirements for the two facilities.

Caution is advised in choosing the most proximate position. The two facilities may move different amounts because of the same wind pressure; one may be unloaded at ambient temperature or fully ice loaded, while the other may be at maximum operating temperature. One could be at initial sag while the other is at final sag. Examination of Rule 233 of the National Electrical Safety Code is recommended.

Determining the closest approach of conductors in winter icing conditions can be confusing. When one line is across the wind and another is with the wind, one loads up with more ice than the other one. In addition, one line can lose its ice before the other one, due to wind or due to thermal loading from line losses. Ice typically forms very near −1.1°C (30°F). If the temperature is significantly colder, the precipitation tends to bounce off rather than coat. However, after the ice has formed, the temperature can fall, bringing with it more heating load and sometimes lighting loads. Thus it is not unusual for an ice-covered conductor to be heated by line losses back up to 0°C (32°F) and eventually melt its ice off. In the meantime, this conductor will be at its maximum sag of 0°C (32°F) with a full coating of ice. The position of the lower conductor for clearances purposes is at the coldest ambient temperature at which the load on the upper conductor could heat it up to 0°C (32°F). This takes care of the cases where the lower conductor has dropped its ice (or was never fully iced) or the lower conductor has come down and is being replaced (without ice) when the upper conductor is at its maximum sag.

Unless the actual relative icing/heating conditions are known, a practical method of choosing an appropriate mismatch in temperature for the winter icing clearances is to use the temperatures used by the National Electrical Safety Code to determine required strengths as the temperature of the lower conductor without ice when the upper conductor is at 0°C (32°F) with the specified ice loading. They are −9.4°C (15°F) for the medium loading district and −17.8°C (0°F) for the heavy loading district. Detailed examination of Rule 233 of the NESC is recommended.

When the suspended facilities are at their closest proximity, both the vertical clearance and the horizontal clearance should be checked to see which is controlling. The horizontal clearance must be at least 1.5 m (5 ft). Where the voltage potential between them exceeds 129 kV, an additional clearance of 0.4 in/kV is required. The basic vertical clearance when the conductors are closest together is 0.6 m (2 ft), with additional clearances required where higher voltages are involved with communications.

Clearances from Buildings and Other Installations. Wires, conductors, and cables which pass by buildings, signs, supporting structures of a second line, pools, bridges, tanks, etc. are required to have clearances from those structures, when not attached to them, that allow normal use of those facilities. Required clearances are given in the NESC. In general, clearances above portions of structures which are accessible to pedestrians or vehicles are the same or similar to those above ground for the same activity. Where pedestrian access is restricted and the area is normally accessible only to workers, lesser clearances are allowed. Clearances allow normal maintenance of the structures being passed by. The minimum horizontal clearance from energized distribution primary voltage conductors to supporting structures of a second line, lighting support of traffic signal support is 1.5 m (5 ft). The minimum vertical clearance above such supporting structures is 1.37 m (4.5 ft).

The vertical clearance required by the NESC for energized distribution primary voltage conductors above building roofs and projections not accessible to pedestrians is 3.81 m (12.5 ft); this allows workers with hand tools the room to work on the roof. Above signs, this clearance drops to 2.4 m (8.0 ft) for open conductors. Vertical clearances to roofs and balconies accessible to pedestrians and catwalks on signs and tanks cannot be less than 4.1 m (13.5 ft) for the distribution primary voltage conductors. Lesser clearances apply for communication and for supply cables and secondary voltages up to 750 V. Greater clearances apply for higher voltages. Bridges have lesser clearances where workers are allowed and generally the same clearances as buildings where pedestrians are allowed. Greater clearances are required around pools and associated structures to accommodate pool skimmer poles, rescue poles, and diving.

Clearances for Facilities Suspended from the Same Structure. The clearances required by the NESC, between wires, conductors, and cables carried on the same supporting structure provide adequate room on the structure for workers to operate and maintain the lines, adequate room out in the span for communication workers to work under supply facilities, and adequate separation between suspended facilities to limit contact by them during operation. Because some communication cables on longer spans will "gallop" in high winds, and because such cables may gallop as high as a straight line between their points of attachment at supports, supply conductors above 750 V generally are prohibited from sagging lower than such a straight line.

The vertical clearance at supports for primary-supply conductors above other facilities is required to be 16 in above neutrals and 40 in above communication (more if above 8700 V). For supply conductors at voltages up to 8700 between conductors, the minimum horizontal clearance provided by the NESC, is 12 in. For higher voltages it is required that 0.4 in be added for each 1000 V above 8700. For sags of more than 24 in, clearances greater than 12 in are required and are determined by the following formulas, where S = sag, in. For wires of No. 2 AWG or larger,

$$\text{Clearance} = 0.3\text{ kV} + 8\sqrt{S/12} \quad \text{in} \tag{18-18}$$

For wires smaller than No. 2 AWG, the rule is

$$\text{Clearance} = 0.3\ \text{kV} + 7\sqrt{(S/3)} - 8 \quad \text{in} \tag{18-19}$$

Multiconductor, spacer, and other cabled types of supply-circuit construction are exempt from the above phase-spacing requirements.

Climbing Space. Climbing space must be provided on poles for workers to move up and through facilities to reach each of the facilities on the structure unless it is the unvarying rule that workers do not climb the poles. If workers climb some portions of the structure and not others, for example, communication but not supply, and the remainder are worked from insulated bucket trucks, only the portions climbed must have climbing space. The climbing space may move around the pole to allow access to other parts or to avoid traveling through certain locations, as long as full space is allowed to facilitate the turns. These dimensions are intended to provide a clear climbing space of 24 in when the conductors bounding the space are covered with appropriate temporary protective covering. The vertical dimension is 40 in above and below the conductors (60 in if above 8700 V). The National Electrical Safety Code, ANSI C2-2002, specifies the horizontal climbing-space dimensions given in Table 18-17. Where conductors of the same voltage classification are on the same crossarm, the horizontal dimensions are projected vertically not less than 40 in above and below the limiting conductors. Equipment such as transformers, regulators, capacitors, surge arresters, and switches when located below the conductors should be mounted outside the climbing space.

TABLE 18-17 Horizontal Climbing-Space Dimensions (Voltage-to-Ground)

Supply conductors	Horizontal climbing space, in
0–750 V	24
750 V–15 kV	30
15–28 kV	36
28–38 kV	40
38–50 kV	46

Note: 1 in = 25.4 mm.

Working Space. Working spaces are required by the National Electrical Safety Code, ANSI C2-2000, at each side of the climbing space and extending to the outermost conductor positions. The size of the working space is linked to the vertical clearances required between conductors at the support and the size of the climbing space, both vertical and horizontal.

Clearances Between Supply and Communication Equipment. The National Electrical Safety Code, ANSI C2-2002, generally requires a minimum of 40-in clearance between supply equipment and communication equipment and between the conductors of each to the equipment of the other. This provides for footroom for the supply workers and headroom for the communication workers. Special provisions are made to allow luminaires to be placed between these facilities.

Vertical and Lateral Conductors. Vertical and lateral conductors may be run within the normal supply space and communication space on a pole if they are so located as to meet the requirements of the National Electrical Safety Code, ANSI C2. Generally such conductors are required to be insulated and placed out of the climbing footroom area or held away from the pole on the opposite side.

18.17 LINE CONDUCTORS

Conductor Factors. Copper and aluminum are the metals most used as conductors in distribution systems. Proportions are fixed by the combined effect of conductivity, weight, strength, and cost. Recent years have seen such a shift in availability and cost that aluminum has gained almost universal use in distribution, supplanting copper, which was preferred for many years.

Conductor Materials. Aluminum has the advantage of about 70% less weight for a given size, but its conductivity is only about 61% that of annealed copper. For distribution, it is commonly rated as equivalent to a copper conductor two AWG sizes smaller, which has almost identical resistance. Its tensile strength is less than copper, and it is commonly used, particularly in the smaller sizes, by stranding aluminum around a steel core of proper size to give the desired tensile strength. In larger sizes, the tensile-strength requirements of distribution are satisfied by stranded aluminum without the reinforcing steel. Another way of obtaining high tensile strength is to combine steel with copper or aluminum wires. Steel is combined with copper in a high-strength strand known as Copperweld, which has 30% to 40% of the conductivity of a copper conductor of equal size. In a similar manner aluminum and steel conductor can be combined into what is known as Alumoweld.

Both copper and aluminum are suitable for use as substation buses, being available in flat bars, tubes, and rods. For very heavy currents, channel shapes are used to make up box-type buses, which are the most economical for such applications.

Use of Copper. Where copper is used for overhead circuits with span lengths of 200 ft or more, it is commonly used in the hard-drawn form because of its greater tensile strength. For common types of local distribution circuit where spans are shorter and flexibility is desirable, medium-hard-drawn, or annealed, copper is used. Mechanical connectors are extensively used for joints and taps on overhead copper.

Underground copper cables are usually made of standard soft copper because of its greater flexibility. The smaller size of copper conductors helps to offset unfavorable price levels because of savings in insulating and sheathing material as well as the ability to put maximum carrying capability in a given size of duct.

Use of Aluminum. In rural line work, where long spans and conductors of high tensile strength are an economic necessity, the combined requirements of conductivity and strength have been met with aluminum stranded around a steel core sized to give the required strength. Such a cable is known as *aluminum cable steel-reinforced* and is commonly designated as ACSR. Development of high-strength aluminum alloys has led to such alternative cables as aluminum conductor alloy-reinforced (ACAR) and all-aluminum-alloy conductor (AAAC), which also combine conductivity with tensile strength. Urban distribution uses ACSR and all-aluminum conductors. Stranded aluminum is common where large conductors are required.

Underground Aluminum Cables. The development of such synthetic insulations as polyethylene has made aluminum almost universally used for underground distribution. In the smaller sizes for URD, a solid conductor is often applied rather than stranded construction. *Jointing of aluminum* requires special care to secure good contact and to guard against corrosion. Jointing is often done with compression devices, although mechanical connectors packed with corrosion-inhibiting compound can be used.

Use of Steel. Steel conductors are rarely used for distribution circuits because of their high resistance. But steel with a heavy covering of copper, known as Copperweld,* or with a heavy covering of aluminum, known as Alumoweld,* has conductivity approaching 40% that of copper and can be used in some applications. Such coated conductors are also very attractive as high-strength strands or reinforcements for composite cables, which get improved conductivity from strands of hard-drawn copper over the Copperweld or hard-drawn aluminum over the Alumoweld.

Conductors reinforced with steel have impedances which increase somewhat as current density increases. Voltage drops are correspondingly higher than those of copper or aluminum conductors of equal conductivity.

Copperweld and Alumoweld are generally more durable than galvanized-steel cables. They have therefore been used to some extent for guy cables. They are also used widely for shield wires.

*Copperweld and Alumoweld are registered trademarks of the Copperweld Bimetallic Group.

18.18 OPEN-WIRE LINES

Crossarms. Southern pine and Douglas fir are the best woods for crossarms because of their thin, straight grain, high tensile strength, and durability. Experience indicates that a cross section 3½ in wide by 4½ in high is ample for the average distribution line. Main lines are commonly built with six-pin arms, and smaller lines use four-pin arms. Minimum spacing of pins is 12 in, and spacings of 14 to 16 in are commonly used. Minimum spacing of pole pins is 30 in to provide climbing space. Crossarms also are used for supporting transformers and other equipment.

Double crossarms are installed on poles at corners, at terminals, and at other points where unusual loads are to be supported. Vertical racks are installed on poles to support secondary and multiple street-lighting wires. They are available with two-, three-, or four-spool insulators. Rack construction is less expensive than crossarms and has supplanted them to a large extent. When services run to houses on both sides of the street, two racks are required, one on each side of the pole. In addition, several pole-top designs mount insulators directly on the pole, eliminating the use of crossarms.

Wire Stringing. In erecting wire, the *tension* should be sufficient to prevent too much sag in the spans and yet not so great as to stress the wire unduly. For practical purposes, the approximate formula given by Rankine may be used:

$$t = S^2 w/8d \qquad \text{lb} \tag{18-20}$$

in which t = tension, lb, S = span length, ft, w = resultant load, including weight of wire, lb/ft of conductor, and d = sag, ft, at the center of a horizontal span. If span length is doubled, tension must be quadrupled in order to keep sag the same. If tension is the same on several spans of different lengths, sag is different in each span. The sag of any span when tension is known is found by changing Eq. (18-20) to the form

$$d = S^2 w/8t$$

Sag Tables. Maximum tension in a span is limited by the strength of the wire and supports. Conductor sags under the assumed loading conditions for the particular loading district (see National Electrical Safety Code, ANSI C2) should be such that the tension of the conductor should not exceed 60% of its ultimate strength. Also, the tension at 60°F, without external load, should not exceed 35% of the conductor ultimate strength under its initial unloaded condition.

It is not unusual to design so that the tension of the conductor will not exceed 50% of its ultimate strength under loaded conditions or a 2000-lb limitation. In some cases, the same sag values are employed for a range of wire sizes so that the appearance of a line carrying different conductor sizes will be improved. The sags given in Table 18-18 are selected from standard sheets of a large utility.

Expansion and Contraction of Conductors with Temperature Change. Changes in sag due to expansion and contraction of conductors under varying temperature conditions are important in the stringing of conductors. Lines erected in winter months are likely to be too slack during the summer unless allowances are made. The length of wire in a span, elastic stretching due to mechanical loading being disregarded, varies as determined by the coefficient of expansion of the conductor and the temperature range,

$$L_t = L_0(1 + \alpha_0 t) \tag{18-21}$$

where α_0 = coefficient of expansion, ft/ft of length/°F above 10°F
t = temperature, °F (above 0°)
L_0 = length of wire at 0°F

For aluminum: α_0 = 0.000024/°C or 0.0000133/°F
For ACSR: α_0 = 0.0000112/°C or 0.0000062/°F (nearly that of steel)
For copper: α_0 = 0.000017/°C or 0.0000094/°F

TABLE 18-18 Sags for Typical Distribution Conductors (Heavy-Loading District—60°F)

Size, AWG or M cmils	Conductor material	Sags (in) for span lengths (ft) of							
		80	100	125	150	175	200	250	300
Open wire:									
No. 1/0	Al alloy, bare		10	16	23	31	40	27*	38*
No. 3/0	Al alloy, bare		10	16	23	31	40	32*	46*
336.4	Aluminum, bare		10	16	23	31	40	75	108
No. 1/0	Al alloy, polyeth.		10	16	23	31	40	59*	90*
No. 3/0	Al alloy, polyeth.		10	16	23	31	40	70	101
336.4	Aluminum, polyeth.		18	27	38	51	66		
Cabled secondaries:									
3 No. 1/0	Al alloy, insul.		10	16	23	31	40		
3 No. 3/0	Al alloy, insul.		18	27	38	51	66		
4 No. 3/0	Al alloy, insul.		18	29	43	60	80		
Cabled service drops:									
3 No. 4	Aluminum, insul.	32	52	73†					
3 No. 1/0	Aluminum, insul.	51	79	116†					
4 No. 3/0	Aluminum, insul.	68	108	171†					

Note: 1 in = 25.4 mm; 1 ft = 0.3048 m; 1 lb = 0.4536 kg.
*Taken up to 2000-lb tension limit for spans over 200 ft.
†For 120-ft service drops, 450-lb limit.

Ampere loadings on distribution circuits often require that the temperature rise of the conductor due to resistance losses must also be taken into account.

18.19 JOINT-LINE CONSTRUCTION

Joint-line construction is used where two or more utilities would otherwise maintain separate pole lines, such as where both power and communication lines are routed along the same street.

Basis of Joint Use. Poles are used jointly under a *joint-ownership agreement* or *under a lease agreement*. Under joint ownership, the cost of providing the pole is borne jointly by the companies which share in its ownership. Division of expense is, in general, made in proportion to the space allotted to respective users. Clearance space, required between power and communication circuits and between the lowest attachment and ground, is disregarded in determining percentage of ownership. Clearance between higher-voltage and lower-voltage power circuits is chargeable to the higher-voltage circuits.

In case of leased space, the lessee acquires only the right to occupy a specified space. The owning company installs and maintains the pole and includes all charges in the rental price. Lessees usually install their own attachments and maintain them, though pin space is sometimes leased where space for only a few wires is required.

Construction Specifications. The types of construction, clearances, and relative levels of different classes of circuit should be provided for by a suitable specification, forming part of the agreement under which joint use is entered into. The general purpose is that construction of all parties be such as not to jeopardize the service or equipment of any of the other parties to the agreement. Construction requirements are set forth in the National Electrical Safety Code, ANSI C2. Some of the most important parts of such a specification are discussed in the following paragraphs.

Relative Levels of Supply and Communication Conductors. When supply and communication conductors are located on the same poles, it is generally desirable that the supply conductors be located at the higher levels. This places the higher voltages near the pole top and the communication conductors at the lower levels. This relative location of facilities provides a lower probability of contact between the two systems since the supply conductors are usually larger than the communication lines. This also provides easier access to the lower-voltage or communication circuits by the line crews and avoids the need to climb through the supply conductors to work on the lower-voltage or communication systems. Where 600-V trolley feeders are carried on joint poles, the feeders are located for convenience at the approximate level of the trolley contact conductor.

Vertical Clearances. Spacing of conductor attachments must be appropriate with the requirements of safety in operation and maintenance. Clearance requirements are spelled out in detail in the National Electrical Safety Code, ANSI C2. Generally, a minimum vertical clearance of 40 in is used between communication conductors (or open wires of 0 to 750 V) and supply conductors operating between 750 and 8700 V to ground. Greater clearances are required for supply conductors operating above 8700 V. If the communication circuits belong to the utility for use in operating supply lines, reduced spacings of 16 and 40 in, respectively, are permitted.

Grades of Construction. Strength of poles must be such as to withstand ice and wind loadings normally experienced in the locality where the line is built, for all of the conductors to be carried. These conditions vary greatly in different parts of the United States, there being no ice in some parts of the country and a greater prevalence of wind in others.

Conductor Size. The size of conductors should be such that they do not experience a tension more than 60% of their rated breaking strength under conditions of maximum loading and not more than 25% of this value for final unloaded tensions at 60°F. Very little use is made in new construction of wire sizes smaller than No. 1/0 stranded aluminum or No. 2 ACSR.

Inductive Coordination. A vast majority of the newer telephone circuits on joint-use distribution lines are in cable, rather than open-wire, and telephone interference is rarely encountered. Where long exposures of open-wire circuits do exist, if may be necessary to make suitable transpositions in both power and communication circuits to eliminate electrical unbalances as much as practicable.

Aerial-Cable Construction. Insulated aerial cables carried by steel messenger cables have been used occasionally in primary distribution circuits where undergrounding is not practicable and special conditions, such as reduced clearances or unusually severe tree problems, exist. This type of cable is fastened to a galvanized-steel cable, or *messenger,* by means of brackets or lashings in a manner similar to large communication cables. Usually it consists of one, two, or three insulated conductors spiraled around the messenger which supports the assembly mechanically and usually serves as a neutral as well. This type of construction is finding very little usage in new system extensions.

Spacer-Cable Construction. Spacer-cable construction provides many of the advantages of aerial cable at lower cost. It consists of primary conductors having less insulation than the cable insulation which is customary for the circuit voltage, supported on a messenger and separated from each other by insulating spacers installed at suitable intervals along the line. Note, however, that the energized phase conductors of spacer cable are not directly cabled together with an effectively grounded neutral. They are held away from the neutral by a plastic spacer. Thus they meet Rule 230D of the National Electrical Safety Code, IEEE C2-1997, and must have the same clearances to other objects as bare conductors.

18.20 UNDERGROUND RESIDENTIAL DISTRIBUTION

During the past 40 years, the evolution of underground distribution systems, particularly single-phase systems to serve residential areas (URD), has proceeded at a rapid rate. For a so-called mature industry, the rate of change has been phenomenal. Costs have been steadily reduced through the introduction of new system concepts, improved installation practices, and the development of specialized equipment.

Nearly every utility in the United States now has a policy covering the installation of URD in new residential tracts. Conditions vary all the way from a substantial differential payment by the developer to a no-charge basis by the utility, although the developer usually is requested to assist with excavation. In addition, a number of states have established legal requirements mandating that all new residential developments in excess of a given number of homes be served by an underground distribution system. As a result, perhaps as many as two-thirds of new residential dwelling units are being served underground.

Cost. Underground distribution systems often cost more than comparable overhead systems. What are the principal factors contributing to the rapid growth of URD? These include

1. Greater public interest in the aesthetic appearance of residential communities.

2. Reduced cost of underground equipment and installations brought about by

Solid dielectric insulated cables—lower-cost—suitable for direct burial without duct systems

Factory-built cable terminations and splices of low cost easily prepared in the field by ordinary line crews without the aid of highly trained cable splicers

Mass production of specialized equipment such as pad-mounted transformers and accessories

Improved installation technique and equipment

Performance. Most observers are of the opinion that the frequency of faults is lower on underground systems than on overhead systems and that the faults are not so likely to "bunch up" because of storm conditions. However, faults are much more difficult and time-consuming to find, to isolate, and to repair on underground systems. This, coupled with the fact that many operating procedures cannot be performed on an underground system while it is energized and that it is impossible to make many of the temporary improvisations on underground circuits that can be accomplished on overhead systems, has led to the development of protective and sectionalizing equipment such as switches and separable cable connectors which often are physically integrated as accessory devices in the underground distribution transformers. In addition, several utilities with significant amounts of older cable have found that they are having more unexpected reliability problems as cables have aged.

Service-restoration requirements also have resulted in primary-system designs which operate as a normally open loop as shown in Fig. 18-41. In the case of a cable fault, this facilitates location and isolation of the failure and more rapid service restoration to all customers on the unfaulted portion of the primary loop. It is estimated that about 85% of primary URD systems are being operated as loops, the remainder being radial. Where radial laterals are used, many utilities provide portable, aboveground cables so that faulted cables can be bypassed temporarily and service restored while repairs are being made.

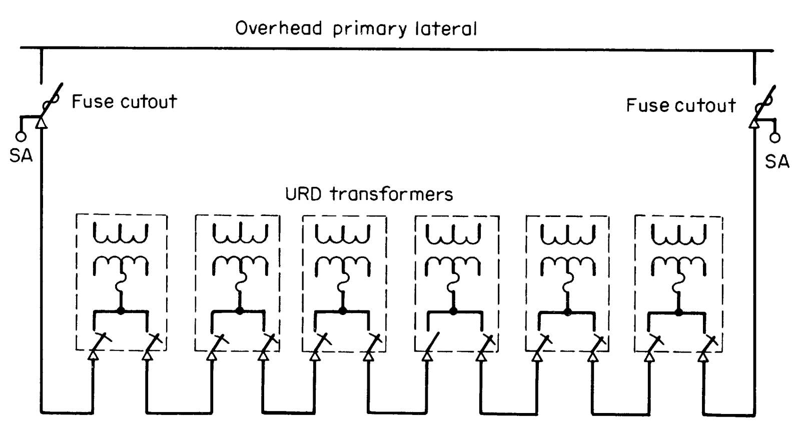
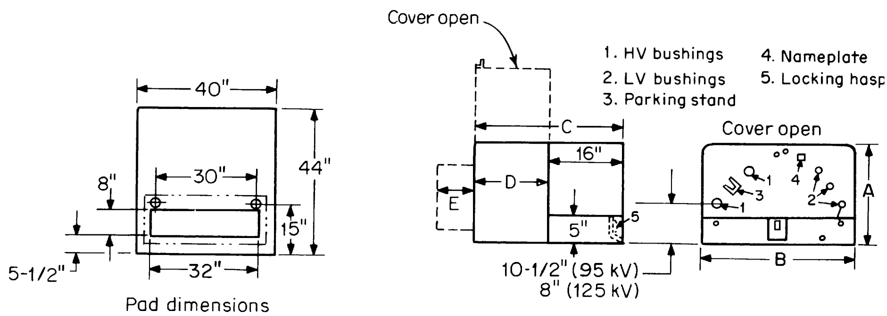


FIGURE 18-41 URD system derived from existing overhead circuits.



95- and 125-kV BIL (35 kV GrdY and below)

kVA	Dimensions, in				
	A	B	C	D	E
15	24	36	33.9	17.9	
25	24	36	33.9	17.9	
37.5	24	36	33.9	17.9	
50	24	36	33.9	17.9	8.0
75	24	36	40.0	24.0	8.0
100	26	36	40.0	24.0	8.0
167	32.5	36	40.0	24.0	8.0

FIGURE 18-42 Mini-pad distribution transformer. (*General Electric Company*.)

Transformers. The heart of the URD system is the single-phase distribution transformer because the primary cable terminations, switching and sectionalizing equipment, and overcurrent protective equipment usually are housed in the transformer enclosure. Thus most operating procedures require access to one or more distribution transformers. Three general types of single-phase transformers are in use.

Pad-Mounted. Figure 18-42 shows the predominant type of transformer being used for URD. The transformer shown is called the *mini-pad*. The term *pad* derives from the fact that transformers in this category usually are installed on concrete slabs, or pads.

The electrical functions of URD transformers cover essentially the same range as pole-type units. For reasons of safety, of course, they must be built in tamper-resistant configurations with no exposed electrically energized parts because of the proximity of such transformers to the general public.

The mini-pad in Fig. 18-42 has its cover open. The two primary bushings at the upper left are for use with load-break, separable insulated connectors, or elbows. This results in a "dead-front" configuration which is required to achieve the low-height mini-pad construction. The three 120/240-V bushings are at the right-hand side.

Many other combinations of pad-mounted construction and accessory equipment are available, including "live-front" primary connections with stress cones for the cables, internal or external primary fuses and switches, secondary circuit breakers, etc. Refer to appropriate product bulletins of the manufacturers or handbooks for further equipment details. Generally, the loadability of pad-mounted transformers is comparable with that of pole types.

Residential Subsurface Transformers (RST). Although usage of pad-mounted transformers predominates, a number of residential subsurface transformers are used. The RSTs are installed in relatively tight-fitting vaults with the cover grating of the vault at ground level.

Cooling is accomplished by natural convection of the air, although some users increase the efficiency of circulation by means of special chimneys to direct and control the circulation. With properly designed and installed chimneys, the loadability of RSTs is equal to that of pole types.

The RSTs must be submersible and therefore utilize dead-front primary cable terminations, usually the separable insulated connectors or "elbows." Provisions for operation of accessories such as switches, fuses, and circuit breakers are located on the cover of the transformer so that they can be operated by a member of the line crew standing on the surface of the ground. Usually the vault is too small for a person to enter.

Primary Cables. Primary URD cables are almost universally of the single-conductor concentric-neutral type employing polyethylene or cross-linked polyethylene insulation. Specifically, the use of TRXLPE and EPR are increasing in usage. Ordinary polyethylene is a thermoplastic which melts at temperatures in the order of 110°C. The process of "cross-linking" polyethylene converts it into a thermosetting material which does not have a melting point, per se.

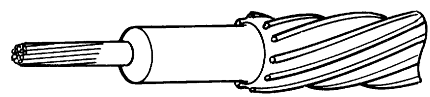


FIGURE 18-43 Concentric-neutral type of primary URD cable.

Figure 18-43 shows a section of primary URD cable. The central conductor is the energized phase conductor, and the external concentric wires serve as the neutral.

Corrosion of the copper concentric-neutral wires of primary URD cables results in reduced cross-sectional area of the wires, increasing their resistance. In some instances, the continuity of the wires is destroyed. Neutral corrosion may cause safety and operating problems on the URD circuit. Corrosion occurs when the neutral wires become anodic, which results in loss of metal. The wires may become anodic due to nearby dissimilar metals or to variations in soil characteristics along the cable route. Determining the location and extent of corrosion damage is a complex procedure which may involve surveys, testing, and in some cases, excavation. Corrective actions for existing cables include replacing portions of the cable, reestablishing the neutral circuit, and installing sacrificial anodes for cathodic protection. Corrosion in new installations can be controlled by the proper selection of materials, cable construction, type of installation, and cathodic protection. The use of jacketed concentric neutral cable to reduce the problem has been increasing over the years and was used by over 80% of respondents in *Transmission and Distribution*'s 1990 survey on underground distribution practices. Most utilities directly bury the primary cables, although the trend to conduit is increasing. Often the URD cables are placed in the same trench as the telephone cables.

The precise calculation of voltage drop in direct-buried, concentric-neutral primary cables is quite complex because a portion of the single-phase load current flows in the concentric-neutral conductors and a portion in the earth surrounding the cable. Also, there may be an induced circulating current in the neutral conductors. Typical values of voltage drop per 100,000 A · ft are shown in Table 18-19.

TABLE 18-19 Single-Phase Voltage Drops per 100,000 A · Ft* for 15- and 35-kV Direct-Buried Concentric-Neutral Cables (Loop Values)

	Voltage class									
	15 kV					35 kV				
	Lagging power factor									
Conductor size	0.7	0.8	0.9	0.95	1.00	0.7	0.8	0.9	0.95	1.00
Underground primary										
Aluminum:										
Concentric-neutral—direct-buried, cross-linked polyethylene, conductor 70°C, neutral 60°C, earth resistivity 90 Ω-cm³, full insulation										
No. 2	44.1	46.6	48.1	48.1	44.8					
1/0	30.9	32.8	34.0	34.1	32.0	31.3	33.0	34.1	34.1	31.7
2/0	25.3	27.0	28.1	28.3	26.8	25.8	27.3	28.3	28.4	26.6
4/0	17.0	18.2	19.1	19.3	18.5	17.5	18.6	19.3	19.5	18.4

*Values in the table give the difference in absolute value between sending-end and receiving-end line-to-neutral voltages, in volts.

To use the table, calculate the ampere-feet as the product of the current in the phase conductor and the distance in feet between the source and the load. The effects of direct burial on impedance of the return current path are included in the tabulated voltage drops.

Secondary Cables. Usually three polyethylene-insulated, single-conductor cables are used for the 120/240-V secondaries and services. These may be separate cables or of triplex construction. In some cases a bare copper neutral conductor is used. The secondary and service cables are usually directly buried.

Cable Terminations. A major advantage of the polyethylene-insulated primary cable, in addition to low cost, is the ease with which it can be terminated in contrast with earlier traditional paper and lead cable, that is, cable insulated with oil-saturated paper with an outer lead sheath or jacket. Termination and splicing of a PILC cable requires a skilled cable splicer working with paper tapes and equipment for soldering and "wiping" the lead cover to potheads or to another section of PILC cable. Several hours are needed to prepare a 15-kV PILC cable termination.

The XLPE concentric-neutral cable can easily be prepared for termination by means of either a factory-made stress cone or a separable insulated connector. This preparation can be done by an ordinary lineworker in a hour or less using cutting jigs and tools to prepare the cable for the installation of the factory-made termination.

When the URD primary cable is terminated by means of a simple stress cone, the electrical connection to the terminal of the connected device usually is an exposed or "live-front" connection. When a separable insulated connection is used for termination, a "dead-front" construction is obtained; that is, all exposed surfaces of the cable and its termination are essentially at ground potential and thus present less of a hazard to operating personnel. In some cases, dead-front configuration allows a reduction in dimensions of the equipment.

Insulated connector modules are available in a great variety of configurations such as the elbow and bushing, multitaps, stand-off bushings for temporary use on parking stands, load-break modules, and T taps. Figure 18-44 illustrates a cutaway view of a switch (load-break) module, and Fig. 18-45 is a similar illustration of an elbow connector and module.

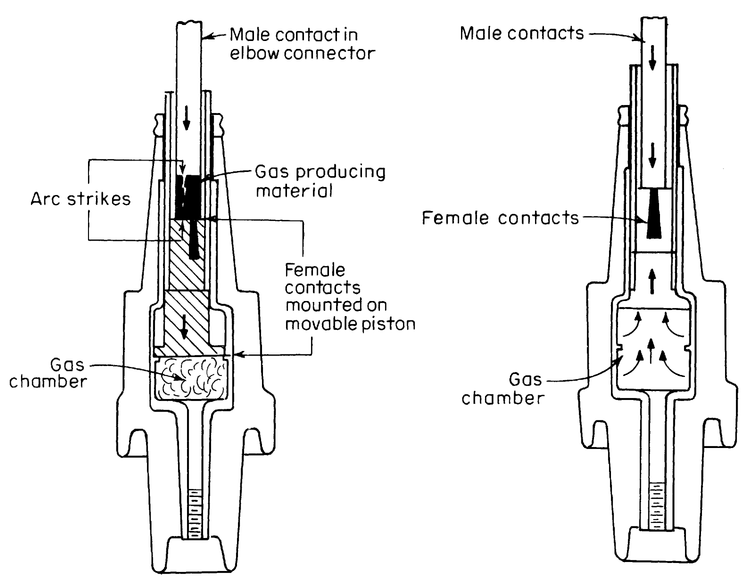


FIGURE 18-44 Piston-action 25-kV connector. (*General Electric Company.*)

The basic separable connector system used with URD distribution transformers usually is rated 200 A continuous and is available to either load-break or non-load-break form. See appropriate product bulletins or handbooks of the manufacturers for further product description.

System Types. About 65% of URD systems are installed along the streets in front of the houses, or "front-lot." The remaining 35% are "rear-lot" systems. There are obvious operating and maintenance problems associated with access to the rear-lot location. At the moment there does not appear to be any strong trend toward either option.

An extremely large number of combinations of transformer equipment are being used, and a detailed discussion of them is beyond the scope of this handbook. However, the following listing is reasonably representative of "typical" practice:

1. Pad-mounted transformer of the mini-pad configuration
2. Primary laterals operated as normally open loops
3. 12.47 grounded Y/7.2-kV primary voltage
4. The primary lateral loops through each distribution transformer, that is, there are two primary cable connections to each transformer (see Fig. 18-41)
5. Front-lot construction
6. Four to eight homes served by each transformer
7. Transformers of dead-front construction load-breaking separable insulated connectors
8. Internal fusing for each transformer
9. Direct-buried, cross-linked polyethylene insulated cables

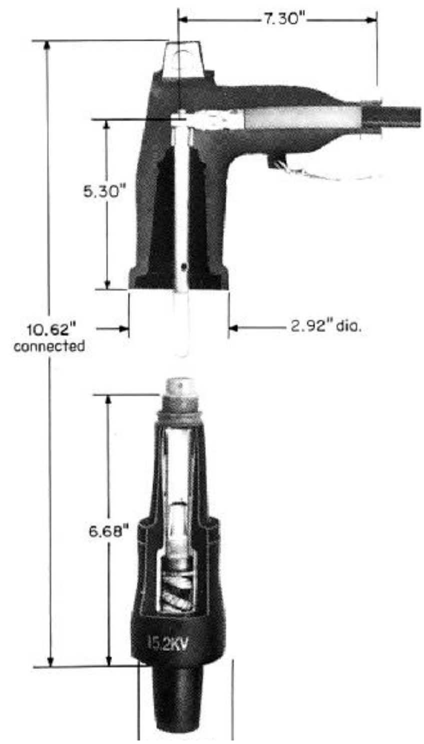


FIGURE 18-45 Cutaway view of elbow connector and switch module. (*General Electric Company.*)

Homes Served per Transformer. There is an optimal number of homes to serve from each transformer depending on the load per home, size of lots, and type of system to be used. For a given load per home and lot size, the cost per kVA of transformer decreases as the number of homes increases. This is so because increasingly larger transformers would be used.

However, as the number of homes per transformer increases, the cost of the secondary and service system increases because of the larger secondary cable required. Since the total cost is the sum of those costs, an optimum number of homes per transformer will exist.

In making such an economic study, it is necessary to examine the secondary-service-system voltage drop. A detailed study also should evaluate transformer and cable losses for the various arrangements.

Four to eight homes served per transformer seems to be reasonably typical of present practice. Larger loads per home and larger lot sizes favor a smaller number of homes per transformer. Conversely, smaller loads per home and smaller lot sizes favor serving more homes from each transformer.

18.21 UNDERGROUND SERVICE TO LARGE COMMERCIAL LOADS

Large commercial loads constitute one of the major segments of utility distribution systems, especially in built-up areas where underground supply systems are a requirement. Demands range from a few hundred to many thousands of kVA per customer, and the engineering and design time to provide adequate service facilities is substantial. Each job is special, requiring selection of appropriate and correctly sized equipment, negotiation of space and layout with building owners or their consultants, and frequently a detailed discussion of facilities, charges, rates, and contracts. The best tool the distribution engineer has is an adequate knowledge of the systems and components which are available, together with guidelines on their cost and reliability. Beyond this, engineering common sense and reasonable operating practices must be combined with the other factors in order to arrive at a decision on the method of service.

Characteristics of Large Commercial Loads. All large commercial loads generally involve the following factors:

1. *Loads* are in the range of 300 to 4000 kVA or more. The larger loads (even up to values of 50 or 75 MVA) are normally supplied by multiples of lower-capacity services.
2. *Utilization voltage* is 480Y/277, although smaller loads may be 208Y/120 and some of the larger institutional loads may be 4160Y/2400 (with the customer providing further step-down).
3. Individual *service size* is limited to about 4000 A by availability of service entrance switching, maximum fuse or breaker sizes, largest commercial wiring systems, and a growing "gut feeling" that this represents enough eggs in any one basket. Providing adequate interrupting capacity is also a definite factor, and single transformers above 4000 A may be priced as specials.
4. *Installation space* is limited and has a high value to the owner. Utility equipment must be as compact as possible and should not require exceptional customer requirements for auxiliaries.
5. Each job is one-of-a-kind and requires much *custom engineering* as well as detailed coordination with the owner of the building facilities. Complex commercial considerations are also involved, covering rates, ownership of facilities, contracts, and future maintenance responsibilities.
6. Service quality must be high, as to both voltage regulation and continuity. Frequent interruptions are not tolerable, and long planned interruptions are not feasible. Service complaints when expressed are long and loud.

Service Arrangements. Several basic service arrangements can be considered for these loads:

1. Radial
2. Primary loop
3. Primary selective
4. Secondary selective
5. Spot network

If radial service were adequate, there would be no need for the succeeding systems because the radial system is the least complex and the least expensive. Unfortunately, when the supply system is underground, it also is the least reliable and generally is unsatisfactory except in special cases. The principal drawback of the radial system is its exposure to long interruptions due to component failure and the necessity for repeated planned interruptions for routine maintenance or new construction.

These five basic service systems are illustrated in Fig. 18-46, which also shows a basic main feeder system of two similar underground feeders.

Radial System. The radial system is exposed to many interruption possibilities, the most important of which are those due to primary cable failure or transformer failure. Either event will be accompanied by a long interruption, reported nominally by utilities as 10 to 12 h. Both components have finite failure rates, and such interruptions are expected and statistically predictable. The system will be satisfactory *only* if the interruption frequency is very low and if there are ways to operate the system without planned outages.

Primary Loop. A great improvement is obtained by arranging a primary loop, which provides two-way feed at each transformer. In this manner, any section of the primary can be isolated, without

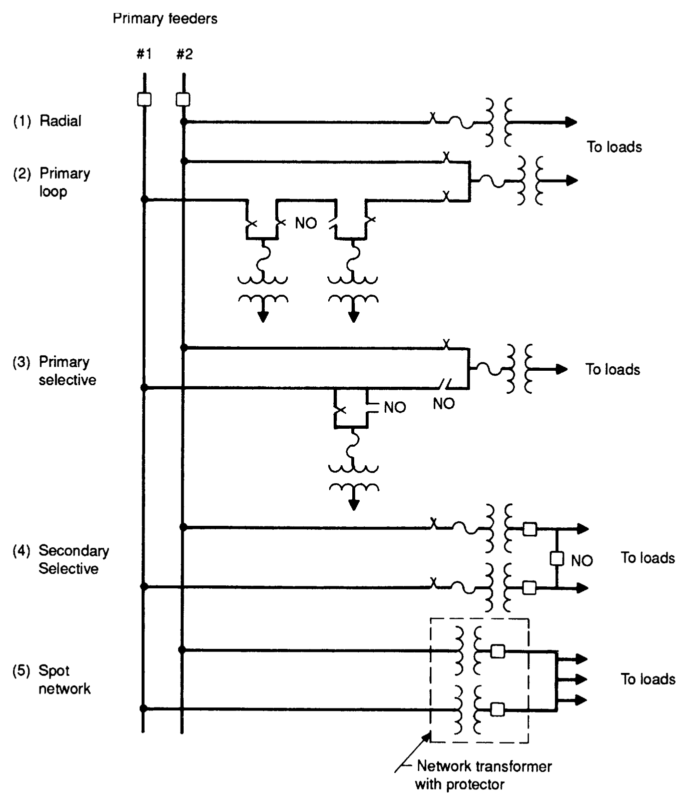


FIGURE 18-46 Five basic service systems.

interruption, and primary faults are reduced in duration to the time required to locate a fault and do the necessary switching to restore service. The cable in each half of the loop must have capacity enough to carry all the load. The additional cable exposure will tend to increase the frequency of faults, but not necessarily the faults per customer. The addition of a loop tie switch at the open point also introduces the possibility of a single equipment fault causing an interruption to both halves of the loop. Murphy's law generally applies to these situations. Automatic loop switching to reduce interruption duration further is very difficult to arrange and is not normally applied to these systems.

Primary Selective. This system uses the same basic components as in the primary loop but arranged in a dual or main/alternate scheme. Each transformer can "select" its source, and automatic switching is frequently used. When automatic, the interruption duration can be limited to 2 to 3 s. Each service represents a potential two-feeder outage (if the open switch fails), but under normal contingencies, service restoration is rapid and there is no need to locate the fault (as with the loop) prior to doing the switching. This scheme is in popular use on many underground systems. Switching times can be improved to less than 1/2 cycle with the use of a static transfer switch (STS).

Secondary Selective. This service system uses two transformers and low-voltage switching. It is not in popular use by utilities for 480-V service but is common in industrial plants and on institutional properties. Primary operational switching is eliminated and with it some causes of difficulty. Duplicate transformers virtually eliminate the possibility of a long interruption due to failure. Load is divided between the two units, and automatic transfer is employed on loss of voltage to either load.

There must be close coordination of utility and customer during planned transfers, and the split responsibility is probably the principal reason for its limited use as a service system.

Secondary Spot Network. Maximum service reliability and operating flexibility are gained by a spot network using two or more transformer/protector units in parallel. The low-voltage bus is continuously energized by all units, and automatic disconnection of any unit is obtained by sensitive reverse power relays in the protector. Maintenance switching of primary feeders can be done without customer interruption or involvement. Spot networks are common in downtown, high-density areas and have been applied frequently in outlying areas for large commercial services where the supply feeders can be made available. This system also represents the most compact and reliable arrangement of components for service in underground systems.

18.22 LOW-VOLTAGE SECONDARY-NETWORK SYSTEMS

Distributed or grid-type secondary network systems have been used for many years by electric utility companies to serve high-density load areas in the downtown section of cities. Secondary networks are used in about 90% of the cities in this country having a population of 100,000 or more and in more than one-third of all cities with population between 25,000 and 100,000.

The service voltage is 208Y/120 V supplying light and power loads in stores, hotels, restaurants, office buildings, apartment houses, and in some cases individual residences. The systems and equipment are entirely underground, and the 208Y/120-V portion consists of grids of interconnected cables supplied at numerous points by network transformers which feed the grid through network protectors.

A given secondary network is supplied by several primary feeders suitably interlaced through the area in order to achieve acceptable loading of the transformers under emergency conditions and to provide a system of extremely high service reliability. Primary voltages are found in the range of 5 to 34.5 kV, with the 15-kV class predominating. See Fig. 18-47.

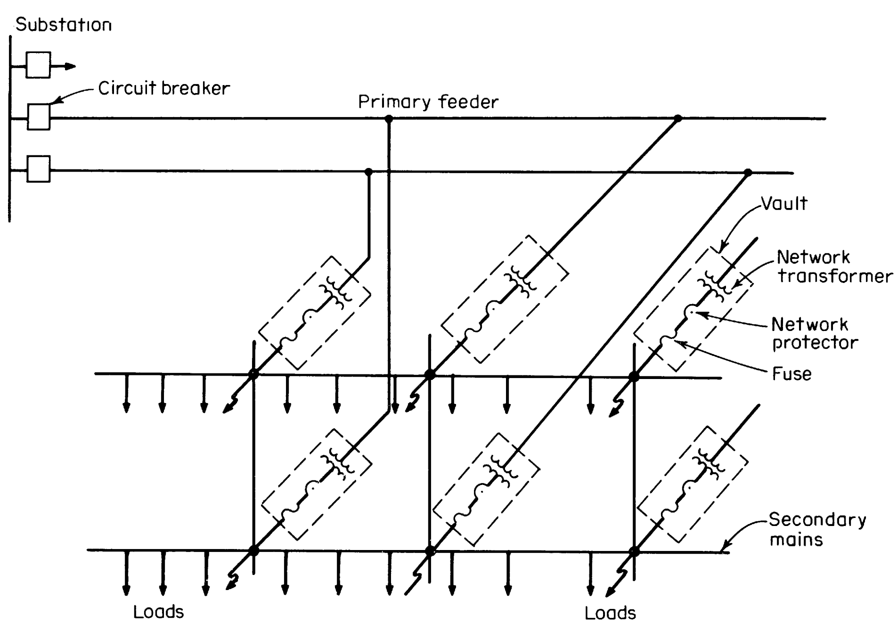


FIGURE 18-47 Schematic diagram of small segment of a secondary network.

The number and routing of the primary feeders are usually based on the assumption that the loss of one or two feeders will not cause a service interruption. For example, the design bogey may be such that the network can operate satisfactorily during the forced outage of one feeder when another feeder is out of service for repairs or maintenance (single contingency).

The secondary cable system is designed so that the loss of one transformer will not cause low voltage or a service interruption. Secondary cable faults in 208Y/120-V networks are allowed to burn clear or are cleared by means of limiters, which essentially are fuses having characteristics proportioned to protect the cable and to coordinate with other protective devices. Secondary faults usually will *not* burn clear on 480Y/277-V networks, so limiters are used extensively in these systems. Usually the secondary mains consist of two or more cables in parallel so that failure of one cable does not result in a service interruption.

Network Transformers. The network protector is mounted on one end of a network transformer, and the primary disconnecting and grounding switch is usually on the other end. In some installations the network protector is isolated from the network transformer, with the connection between the low-voltage terminals of the transformer and the protector made with insulated low-voltage cables. The typical network transformer is 3-phase, 216Y/125 V, oil-cooled, in a heavy corrosion-resistant tank suitable for installation in subsurface vaults under streets or sidewalks. Occasionally, there are installations in dry locations where submersible construction is not required. Five hundred kVA is a very common rating, although 750- and 1000-kVA units are available and in use. (For *spot* networks of 480 Grd Y/277 V, transformer ratings are available through 2500 kVA.) There are only two or three distributed street networks at 480Y/277 V in the United States.

Network Protectors. The network transformer is connected to the secondary network through a network protector (NWP) as shown in Fig. 18-47. The NWP is an air circuit breaker with relays and auxiliary devices and backup fuses, all enclosed in a metal case, which usually is physically mounted on the secondary side of the transformer. The functions of the relays are

1. To open the NWP on power-flow reversal, or in case of a fault in the transformer or in the primary feeder

2. To reclose the NWP when the voltage of the primary feeder is of the correct magnitude and phase relation with respect to the network voltage so that when the NWP closes, power (watts) and vars will flow *from* the feeder *into* the network

Thus, if there is a fault on a primary feeder, it is cleared by operation of the feeder breaker at the substation and the opening of all network protectors on transformers supplied by that feeder. Also, if a feeder breaker is opened manually in preparation for maintenance work on the feeder, all NWPs on that feeder should open automatically because of the reverse power flow caused by excitation of the transformers from the low-voltage side.

Cables. All primary and secondary cables are routed along the streets in duct lines as indicated in Fig. 18-47. Loads are served along the streets and at intersections as shown. Primary cables traditionally have been paper-insulated, lead-covered (PILC), but the solid-dielectric insulated cables have gained rapid acceptance. Secondary cables have commonly used rubber insulating materials, but polyethylene and ethylene propylene rubber insulations now are used extensively. Manholes at street intersections are large enough to hold numerous cable connections and limiters and to allow workers to pull and splice cables.

Continuity of Service. Continuity of service is the outstanding advantage of a network system. When a failure occurs in a primary feeder or in a transformer, the faulty feeder is automatically disconnected, and service continues without interruption. Secondary cable faults are allowed to burn clear or are cleared by means of limiters without loss of service. Substations supplying networks are so designed that typical substation faults will not shut down the network; this is further enhanced by careful interlacing of the primary feeders through the load area and their connection to different bus sections in the substation. It is strongly recommended that a given secondary network be supplied from only one substation. If a network is supplied, for instance, from two different substations, it is possible under some system conditions for power to flow from one substation to the other through the secondary grid and network transformers. Should this occur, some network protectors could "see" reverse power flow and open, thus resulting in the undesirable disconnection of these transformers from the network.

Network Size. A secondary network supplied by five or more feeders will keep transformer loadings at 125% or less during the outage of one primary feeder. If the feeders are in the 15-kV class, each feeder could easily supply six 1000-kVA or twelve 500-kVA network units. Thus five feeders interlaced could supply a 30,000-kVA network under idealized conditions. With 500-ft-square blocks, 1000 kVA per block corresponds to 112 MVA of load per square mile. Some utilities plan for the emergency outage of one feeder while a second is out of service for maintenance.

In general, 208Y/120-V networks are in the order of 30,000 to 40,000 kVA in size. There are many networks smaller than this range and some larger. One important limitation to the size of a secondary network is the ability to restore service after the network has been shut down.

Spot Networks. New commercial buildings in existing 208Y/120-V network areas usually have very large electric loads. These loads frequently are supplied by 480Y/277-V *spot* networks, since it is impractical to handle individual loads much larger than about 200 kVA from the 208Y/120-V street networks, and 480Y/277 V is an excellent voltage for supplying large commercial buildings. In some cases the spot networks are supplied from primary feeders which also serve a distributed network. Early 480Y/277-V spot networks used the same overcurrent protective devices employed successfully for clearing faults in the 208Y/120-V networks. Included are the network relays in the protector for detecting faults on the primary feeders, and the network protector fuses, cable limiters, and service fuses for detecting and isolating faults in the secondary systems. Some 208Y/120-V systems do not use cable limiters, as faults in the 208Y/120-volt network systems are self-clearing under some circumstances.

Operating experience with the 480Y/277-V spot network systems showed that some faults were arcing in nature, drawing significantly less current than that available for a bolted fault. Many of these faults would not self-clear or would self-clear only after extensive damage was done at and around the original point of fault. Also, some of these faults did not draw sufficient current to blow fuses, or else blew fuses only after significant damage was done at the point of fault.

As a result, some utilities have installed devices for detecting and clearing low-current arcing faults in the 480-V portions of the spot-network system. Heat sensors and ground-fault relays are the most commonly installed devices for the detection function. Clearing has been accomplished by tripping of the network protector, which is effective only for faults downstream from the protector terminals. In a few instances, vacuum circuit breakers or interrupters have been installed on the high-voltage side of the network transformer to clear faults in the associated network transformer, network protector, and other portions of the 480-V system.

Network Monitoring. Remote monitoring capability has been added to a few in-service secondary network systems to automatically gather data needed for the operation and planning of the system. Heretofore, such data were obtained from manual measurements at vaults and manholes. With remote monitoring, the data are continuously collected and transmitted to an operations center or other location for manual and computer analysis. Telephone lines, power-line carrier, two-way radio, and fiberoptic cable have been used as the communications links from the network vaults and other monitored sites of the substation or elsewhere. Virtually any quantity which can be digitized can be monitored. Examples of monitored quantities are the load on the protectors, protector position (open or closed), network protector fuses status (okay or blown), network transformer temperature, vault temperature, bus voltage, and vault water level. In several monitoring systems with two-way communications, it is possible to remotely trip or close the network protectors.

High-Rise Buildings. Primary-voltage feeders are being used as the riser feeder in high-rise buildings. A rule of thumb is that if an apartment building is 10 floors or more, it is more economical from an overall point of view to use the primary voltage rather than utilization voltage for the riser feeder. Similarly, for a commercial building with 480Y/277-V utilization, a building of 50 floors or more usually justifies the use of primary voltage feeders as risers in the building.

The primary system pattern within a high-rise building is usually a loop as shown in Fig. 18-48 or multiple as shown in Fig. 18-49. This will allow cable faults to be manually isolated by proper switching so that cable faults will cause only short interruptions to customers. Customers fed radially through a transformer will be without service if the transformer fails until a temporary connection can be made to an adjacent transformer. For more important loads like elevators, hall lighting, and fire pumps, better reliability is often obtained by using a spot network or low-voltage selective system.

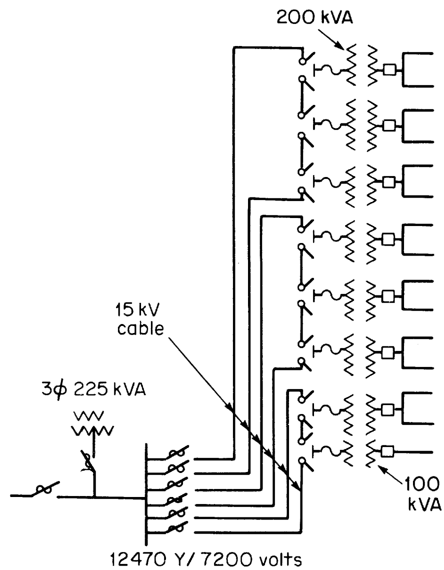


FIGURE 18-48 Schematic diagram of a looped primary system.

Dry-type transformers using air as the insulating medium of the transformers are the most desired for high-rise buildings. This is so because no special provisions have to be made in the transformer room, such as fireproofing for oil-filled or venting the transformer to the outside of the building for nonflammable liquid-filled transformers. Many of the transformer rooms for apartment buildings and some commercial buildings are in the core of the building, which makes it difficult to use the liquid-filled transformer. However, transformers for supplying heavy loads, such as air conditioning in commercial buildings, usually can be located against an outside wall on machinery floors of the building. This makes it relatively easy to vent a nonflammable liquid-filled transformer to the outside of the building. Hence network transformers are often used for this application.

The primary load-break switch or load-break connector with a fuse can be arranged for either the multiple or loop type of feed where the short-circuit current available is within their rating. Usually the primary short-circuit current available is in the 8000- to 10,000-A range. The current-limiting type of fuse is often used for this "inside-the-building" application because it does not discharge ionized gases or noise during interruption.

18.23 CONSTRUCTION OF UNDERGROUND SYSTEMS FOR DOWNTOWN AREAS

Underground construction is required in the built-up downtown areas of cities where the distribution system serves a multiplicity of concentrated commercial loads. Usually the distribution circuits are installed in conduits or duct lines along the city streets. Equipment such as switches and transformers is installed in vaults under the streets or sidewalk or in rooms located within the buildings.

Inflexible conduit systems have not been used widely in the United States except for the early Edison systems and are now completely obsolete. In this system, the conductors were insulated copper rods which were placed in an iron pipe which was then filled with a melted asphaltic compound which solidified on cooling. The tube sections were laid in a trench, joined together, and directly buried. For many years, inflexible conduit systems also were used in Europe; however, in the United States flexible systems gained preference because of the expense and inconvenience of digging through street surfaces in order to make repairs or replacements.

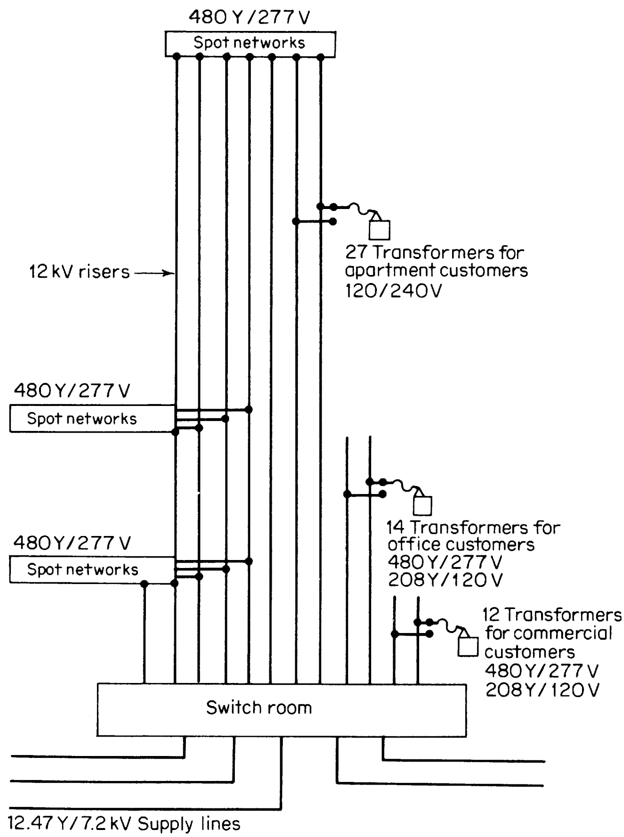


FIGURE 18-49 One-line diagram of the multiple primary system for John Hancock Center.

The flexible undergound system consists of ducts or pipes extending between manholes. This type of system has the advantage of minimum disturbance of street pavement and interference with traffic. Cables can be drawn or withdrawn from manhole locations for repairs or changes. Manholes are placed at all junction points, corners, and as needed for secondary and service cables. The spacing of manholes depends on the types of circuits installed, varying considerably between through-type and local distribution circuits. In straight runs, the intervals may be as great as 500 to 700 ft, depending on the allowable cable-pulling tensions and utility practice.

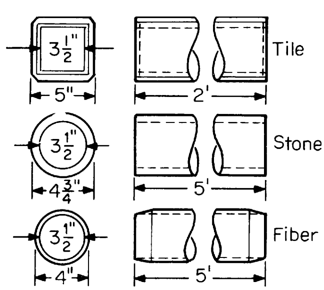


FIGURE 18-50 Types of single duct for underground conduits.

Duct Materials and Practices. Many types of suitable materials have been used for cable ducts, such as fiber-clay tile, concrete, plastic, fiberglass, and soapstone. Preference varies from one utility to another. In general the material should be impervious to water and not degraded by chemical action or electrolysis. The bore usually is round and should be smooth to avoid damage to the cable sheath or jacket. The diameter of the bore should be adequate to accept the largest-diameter cable envisioned for the foreseeable future. Several types of single duct are shown in Fig. 18-50. Diameters ranging from 3½ to 6 in are common.

For underground distribution systems, a duct line usually is built up of a number of single ducts, often arranged in a rectangular array and encased in concrete as shown in Fig. 18-51. For secondary-network systems, duct lines containing 6 to 12 ducts are frequently used.

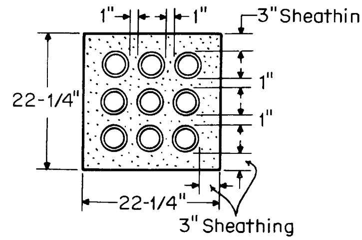


FIGURE 18-51 Arrangement of conduit with concrete sheath.

Number of Ducts. The number of ducts in a given duct line should be sufficient to accommodate anticipated load growth for a number of years in the future. Theoretically, the most economic form of duct structure would be two ducts wide. However, when more than six or eight ducts are required, this design may lead to excessive depth. As a result, usually a rectangular construction, three or four ducts wide and three or four ducts deep, is used, as shown in Fig. 18-51. The maximum number of ducts to be put into a duct line is governed chiefly by thermal limitations. It is desirable to have as many ducts as possible on the outside of the bank in order to facilitate heat transfer to the surrounding earth. Insofar as possible, the inner ducts should be used for cables which produce little heat. Since space for training cables in manholes is limited, it becomes more difficult to rack and train cables when the incoming duct line is several ducts wide and several ducts deep.

Location of Manholes. Manholes provide protected and accessible space in which cables and associated equipment can be operated efficiently. They must be provided in sufficient number to permit pulling in cable without excessive tension, to house necessary transformers and switching equipment, and to provide for splices and service connections.

On long runs the manhole spacing usually is not more than 500 or 600 ft. Where local distribution circuits are involved with numerous service connections, manholes may be located about 100 ft apart. Manholes or vaults to house transformers must be large enough to provide working room and proper ventilation. Location of transformer vaults in the sidewalk area is preferred, and sidewalk gratings are commonly provided to improve ventilation. For locations in the street or in areas accessible to vehicular traffic, the vault roof and gratings must be designed to withstand anticipated loadings.

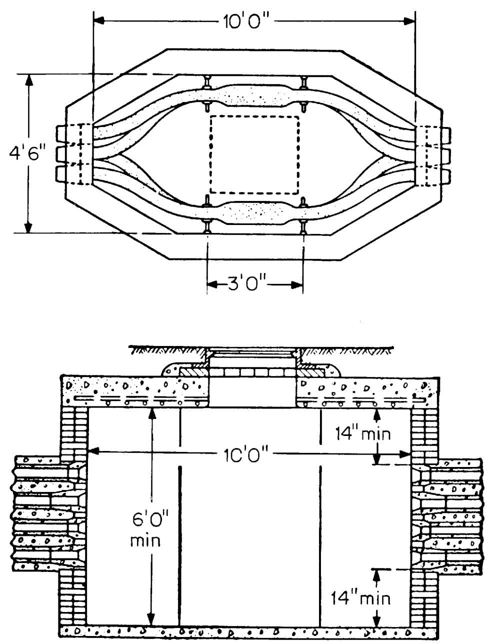


FIGURE 18-52 Straight-type manhole.

Many sizes and shapes of manholes are used. The design used for a particular installation may well be governed by the presence of local obstructions such as gas lines, water pipes, or conduit lines of other utilities. For cable manholes in the streets, many utilities have standardized on a rectangular or coffin-shaped design with the long axis parallel to the duct line. The manhole should be deep enough to allow the lowest duct to enter about a foot above the floor and should have 5 to 6 ft of clear headroom for workers. Also, the bottom of the manhole should be higher than the adjacent sewer system so that a drain can be effective in keeping it dry. Figures 18-52 and 18-53 illustrate two types of cable manholes.

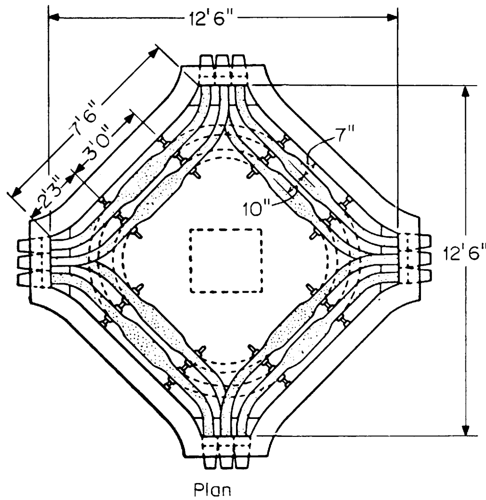


FIGURE 18-53 Nine-duct *X*-type manhole.

Handholes. In some cases, a shallow form known as a *handhole* is used for local-distribution circuit connections. These are usually built above the conduit line so that only the top row of ducts enters the handhole. Secondary distribution circuits are thus accessible for service taps and do not interfere with through lines in the lower ducts.

Installation of Conduit System. Typically there is considerable congestion of underground structures under city streets. Therefore, when a new duct line is planned, a survey should be made to select a location which will present as few obstructions as possible. This is done by noting the position of manholes of existing systems and by consulting whatever map records are available.

The exact final depth of the ditch often cannot be determined until the depths of pipes and conduits crossing it have been disclosed by excavation. Alignment and grades should be established with surveying instruments to ensure proper drainage and avoid pockets. The line may be curved slightly to avoid obstructions, but dips should be avoided where water may accumulate and freeze. Where conditions are suitable, the ditch can be dug so that the earth can be used as the form for the bottom and sides of the concrete encasement.

Structural Design of Cable Manholes. The walls of manholes generally are constructed of brick, nonreinforced concrete, or reinforced concrete. The reinforced-concrete manholes may be poured in the field or may be precast. Many variations in detailed structural design are found with different utilities. The roof must have sufficient strength to support the heaviest street traffic passing over it, which often necessitates use of steel reinforcement or structural steel beams.

To provide access of personnel and the installation of equipment, manhole frames and covers are provided and are supported on the roof of the manhole. The covers for the use of personnel usually are round and made of cast iron or steel. While some square or rectangular covers are used, they usually are heavier for the same effective opening and can fall into the manhole during replacement.

Cable supports or hangers usually are mounted on the walls of the manhole to support the cables in their trained position and to maintain an orderly arrangement of the cables.

Transformer Vaults. When transformers or other equipment are installed beneath streets, sidewalks, or alleys, manholes or vaults are provided. Usually enough room is provided around the equipment so that workers can operate or maintain it. In some cases involving nonnetwork service, commercial loads are supplied from “commercial subsurface transformers” where access to accessory equipment, such as fuses, internal switches, and separable cable terminations, is available from ground level; the vault may be close-fitting since it is not necessary for workers to enter the vault.

The most prevalent types of transformer vaults are found in secondary network systems. They may be located under sidewalks or streets or partly or entirely within buildings. The arrangement, size, and shape of a network vault are determined by the number and rating of network transformers to be installed and the nature of accessory equipment, such as primary switches. Figure 18-54 shows the general arrangement of a sidewalk vault arranged for two network units with network protectors and primary switches. The roof consists of removable slabs of sidewalk concrete (not shown), and access is available at either end by iron steps. Under normal conditions, the entrances are covered with suitable metal gratings or grills which allow for circulation of air.

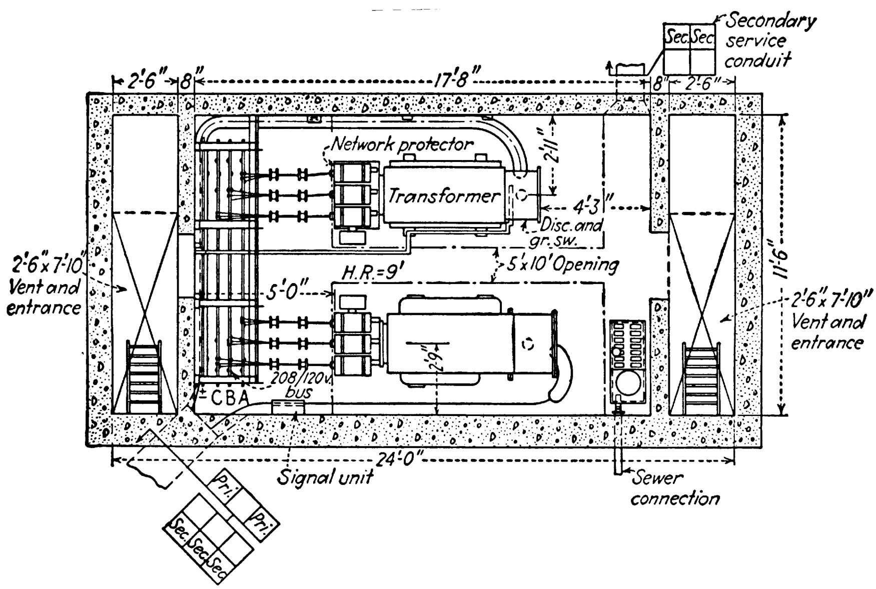


FIGURE 18-54 General arrangement of a network vault under a sidewalk (plan view).

When a nearby sewer is readily available, a drainage connection is often used in street vaults. Sidewalk vaults often do not accumulate enough water to justify a sewer connection and usually are provided with a small sump to facilitate pumping out with portable equipment, if necessary.

18.24 UNDERGROUND CABLES

Types of Cables. Underground distribution systems have been in use for many years in the downtown built-up areas of American cities. In most instances these are secondary network systems with facilities installed beneath streets and sidewalks, and the cables are usually installed in conduit or duct systems. For primary voltage circuits from 5 to 35 kV, paper-insulated, lead-covered (PILC), 3-conductor cable has been used extensively. Single-conductor secondary cables with rubber insulation and neoprene jacket are common. More recently, single-conductor polyethylene-insulated cables are being used for both primary and secondary. Copper conductor predominated in the past, but aluminum has nearly displaced copper in new installations, except where existing duct space is limiting.

In residential and suburban areas, new underground distribution systems to serve commercial loads, such as shopping centers and commercial and industrial parks, often employ direct-buried cables; conduits may be provided in locations where subsequent excavation would be excessively expensive or inconvenient. Aluminum conductors are almost universal. For primary cables, solid-dielectric insulation is used almost exclusively, with cross-linked polyethylene and EPR insulations predominating. Concentric-neutral wires are common. Secondary cables in these systems generally have aluminum conductors and solid-dielectric insulation, with cross-linked polyethylene being the most common. The secondary neutral is usually an insulated conductor, although there is some use of bare copper neutrals.

For most distribution circuits in the 5-kV class or higher, the cables employ a shielded construction. *Shielding* is the use of a conducting or semiconducting material on the surface of insulating material to confine the electric field to the insulation proper and to avoid undesired concentrations of electric stress. Shielding is used on the outer surface of the cable insulation or directly over the main conductor, or both. Outside shielding, often in the form of metallic tapes, metallic sheaths, or concentric wires, must be effectively grounded. This shielding also provides a return path for short-circuit current in the event of cable failure and protects workers from the shock of charging current.

Number of Conductors. Cables can be classified as single-conductor, 2-conductor, 3-conductor, etc., according to the number of separately insulated conductors enclosed by a single sheath or jacket (Fig. 18-55).

Cable Insulation. Electric supply cables are insulated with a wide variety of insulating materials depending on voltage ratings, type of service, installation conditions, etc. In the past, the following have been commonly used:

1. Rubber and rubberlike for 0 to 35 kV
2. Varnished cambric for 0 to 28 kV
3. Impregnated paper of the solid type for voltages up to 69 kV and with pressurized gas or oil up to 345 kV or higher

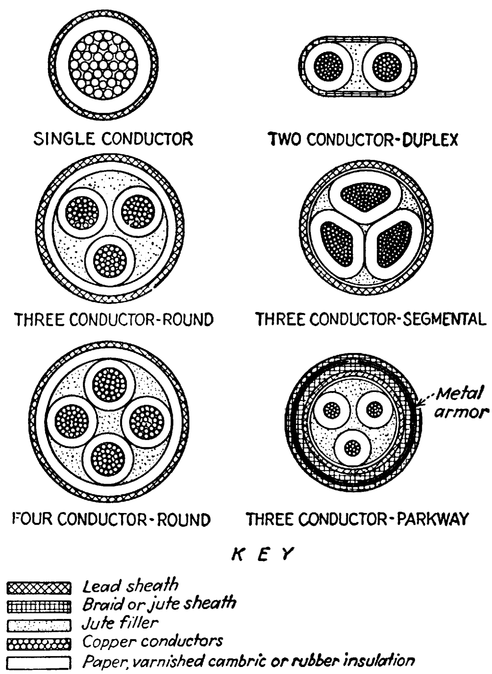


FIGURE 18-55 Cross sections of typical cables.

These insulation systems usually require a sheath or suitable jacket to prevent infiltration of moisture, loss of oil, gas, or impregnant, and to provide protection against corrosion and electrolysis. In some cases, an armor overlay is used to provide mechanical protection. With impregnated-paper insulation of the solid type, a lead sheath is usually provided.

A wide variety of joints, splices, and terminations is used, depending on the voltage, cable insulation, number of conductors, jacketing or sheathing material, and method of shielding. Joints, splices, and terminations are discussed in more detail later in this section.

Single-conductor cables are used, of course, in single-phase primary system for residential service and normally are used in single-phase or 3-phase secondary systems where many taps and connections are involved. Single-conductor cables also are frequently used in direct-buried, 3-phase primary systems. Three-conductor primary cables are often used in duct systems where they have the advantage of occupying only one duct. Several typical cables are shown in Fig. 18-55.

At the present time, solid-dielectric insulating materials such as tree retardant, cross-linked polyethylene, and EPR are receiving the widest application in underground distribution systems, both direct-buried and duct systems. Principal reasons for the wide usage of these insulations are

1. Low cost.
2. Suitability for direct burial or for use in duct systems.
3. Sheath or jacket not generally required.
4. Much easier to tap, splice, and terminate than systems such as solid impregnated paper. Factory-made splices, connectors, and terminations are available and widely used.
5. Excellent mechanical and electrical properties.

In a 2004 survey on procurement practices relating to underground distribution cable covering 60 investor-owned utilities and representing 70% of the total investor-owned utility customers, Dudas and Fletcher reported the following trends:

TABLE 18-20 Insulation Thickness for Cross-Linked, Thermosetting, Polyethylene-Insulated Cable

Rated circuit voltage, phase-to-phase volts	Conductor size, AWG or kcmil	Insulation thickness for 100% and 133% insulation levels	
		mils	mm
0–600	14–9	45	1.14
	18–2	60	1.52
	1–4/0	80	2.03
	225–500	95	2.41
	525–1,000	110	2.79
601–2,000	4–9	60	1.52
	8–2	70	1.78
	1–4/0	90	2.29
	225–500	105	2.67
	525–1,000	120	3.05
2,001–5,000	8–1,000	90	2.29
5,001–8,000	6–1,000	115	2.92
8,001–15,000	2–1,000	175	4.45
15,001–25,000	1–1,000	260	6.60
25,001–28,000	1–1,000	280	7.11
28,001–35,000	1/0–1,000	345	8.76

Note: *100% level* applied where system overcurrent protection is such that ground faults are cleared within 1 min. Applies to the great majority of distribution systems. *133% level* applied where clearing time of 100% level cannot be met, but there is assurance of fault clearing within 1 h. Minimum-size conductors should be in accordance with above values to limit maximum voltage stress on the insulation at the conductor to a safe value.

Source: Adapted from IPCEA Pub. S-66-524, NEMA Pub. WC-7-1471. Revision No. 3, September 1974.

- The survey indicated a significant preference for tree retardant crosslinked polyethylene (TRXLPE) over EPR cable.
- The majority specified concentric neutral on 200 A cable and LC shield, flat strap, flat wire or tape shield on 600 A cable.
- 78% specified an encapsulating jacket rather than an overlaying jacket.
- 95% of the utilities specified an insulating polyethylene compound for their cable jackets.

Insulation thickness for typical cross-linked, polyethylene-insulated, nonjacketed distribution cables are given in Table 18-20. Thickness for most ratings of non-cross-linked polyethylene cables are essentially the same.

Cable Diameters. Overall diameter D of a cable may be computed from the diameter of its conductors d, the thickness of its conductor insulation T, its belt insulation t, and its lead sheath S, as follows:

Single-conductor: $$D = d + 2T + 2S \qquad (18\text{-}22)$$

2-conductor: $$D = 2(d + 2T + t + S) \qquad (18\text{-}23)$$

3-conductor: $$D = 2.155(d + 2T) + 2(t + S) \qquad (18\text{-}24)$$

4-conductor: $$D = 2.414(d + 2T) + 2(t + S) \qquad (18\text{-}25)$$

These formulas apply to conductors of circular cross section. For sector-type 3-conductor cables, the overall diameter

$$D_3 = D - 0.35d \quad \text{approx.} \qquad (18\text{-}26)$$

Electrical Characteristics of Cable. Skin effect is an ac phenomenon whereby alternating current tends to flow more densely near the outer surface of a conductor than near the center. That is, the magnetic-flux linkages of current near the center of the conductor are relatively greater than the linkages of current flowing near the surface of the conductor. The net effect is that the effective resistance of the cable is greater for alternating current than for direct current. This effect increases as the conductor size increases and as the frequency increases. It is also a function of the relative resistance of the conductor material, being less for materials of higher resistance; for example, the skin effect for a given diameter of cable is great if the material is copper rather than aluminum. Because of skin effect, large cables are sometimes built up over a central core of nonconducting material.

The nonuniform distribution of alternating current across the cross section of the cable also has the effect of reducing the effective internal inductance of the cable. Usually, this effect is extremely small in distribution circuits and is neglected.

It should be noted that magnetic flux linking the cable because of nearby current also can affect the cross-sectional distribution of current and can significantly change the effective ac resistance of the cable for multiconductor cables or cables in the same duct. This is known as the *proximity effect*. Most tables of conductor characteristics list factors which combine the results of the skin effect and proximity effect.

If an insulated cable has an outer metallic wrapping such as sheaths, metal pipes, or concentric-neutral conductors installed in such a manner that induced circulating currents can flow normally in these external conductors, losses will occur in these circuits, reducing the ampacity of the cable.

Skin-Effect Coefficients. Skin-effect and proximity-effect coefficients are given in Table 18-20 for copper and aluminum conductors at 25°C. To determine the skin effect on the effective resistance of

TABLE 18-21 DC Resistance and Correction Factors for AC Resistance

Conductor size, AWG or kcmil	DC resistance, Ω/1000 ft at 25°C*		AC resistance multiplier			
			Single-conductor cables†		Multiconductor cables‡	
	Copper	Aluminum	Copper	Aluminum	Copper	Aluminum
8	0.6532	1.071	1.000	1.000	1.00	1.00
6	0.4110	0.6741	1.000	1.000	1.00	1.00
4	0.2584	0.4239	1.000	1.000	1.00	1.00
2	0.1626	0.2666	1.000	1.000	1.01	1.00
1	0.1289	0.2114	1.000	1.000	1.01	1.00
1/0	0.1022	0.1676	1.000	1.000	1.02	1.00
2/0	0.08105	0.1329	1.000	1.001	1.03	1.00
3/0	0.06429	0.1054	1.000	1.001	1.04	1.01
4/0	0.05098	0.08361	1.000	1.001	1.05	1.01
250	0.04315	0.07077	1.005	1.002	1.06	1.02
300	0.03595	0.05897	1.006	1.003	1.07	1.02
350	0.03082	0.05055	1.009	1.004	1.08	1.03
500	0.02157	0.03538	1.018	1.007	1.13	1.06
750	0.01438	0.02359	1.039	1.015	1.21	1.12
1000	0.01079	0.01796	1.067	1.026	1.30	1.19
1500	0.00719	0.01179	1.142	1.058	1.53	1.36
2000	0.00539	0.00885	1.233	1.100	1.82	1.56

Note: 1 ft = 0.3048 m.

*To correct to other temperatures, use the following:

For copper: $R_T = R_{25}\,[(234.5 + T)/259.5]$

For aluminum: $R_T = R_{25}\,[(228 + T)/253]$ where R_T is the new resistance at temperature T(°F) and R_{25} is the tabulated resistance.

†Includes only skin effect (use for cables in separate ducts).

‡Includes skin effect and proximity effect (use for triplex, multiconductor, or cables in the same duct).

a single-conductor 1000-kcmil copper cable operating at 25°C, refer to Table 18-21, where the dc resistance is 0.01079 Ω/1000 ft and the skin-effect coefficient is 1.067. The effective resistance at 60 Hz is 1.067 × 0.01079 = 0.0115 Ω/1000 ft, 6.7% greater than for direct current. For a similar 2000-kcmil, the increase in resistance for alternating current is 23.3%; the ampacity of the cable is reduced to 100/1.233 = 81.1%.

The last two columns of Table 18-21 give the coefficients for multiconductor cables or cables in the same duct. They are used in the same manner as in the previous examples. For the larger conductors, the derating factor is substantial.

Electrostatic Capacitance. The capacitance of a shielded or concentric-neutral single-conductor cable is

$$C = \frac{0.00736K}{10^6 \log_1(D/d)} \tag{18-27}$$

where C = capacitance, farads/1000 ft
K = dielectric constant of insulation
D = diameter over the insulation
d = diameter over the conductor shield

Charging Current. The charging current of a single-conductor cable is

$$I = \frac{0.0463EfK}{1000 \log_1 (D/d)} \tag{18-28}$$

where E = voltage to neutral, kV
f = frequency, Hz
I = amperes per 1000 ft, charging current

For overhead circuits at distribution voltages and power frequencies, the charging current usually is negligible. It may become significant in high-voltage transmission circuits, as discussed in Sec. 14. For insulated cables, the charging current is relatively greater than in overhead circuits because of close spacing and the higher dielectric constant of the cable insulation; $K = 1$ for air and 3.3 for impregnated paper. For unfilled polyethylene $K = 2.3$, and it may run as high as 2.9 for filled, cross-linked polyethylene.

Geometric Factors. Charging current of 3-phase three-core cable is affected by arrangement of conductors (round or sector) and by relative thicknesses of conductor insulation T and belt insulation t. These relations have been put into usable form by working out logarithmic denominators of the equation for various ratios of thickness of insulation to diameter of conductor. This has been termed the *geometric factor*. Charging current of a three-core 3-phase cable is

$$I = \frac{3 \times 0.106EfK}{1000G_2} \quad \text{A/1000 ft} \tag{18-29}$$

For impregnated-paper cable, K is 3.3, and the equation for 60-Hz circuits becomes

$$I = \frac{3 \times 3.3 \times 0.106 \times 60E}{1000G_2} = 0.063\,\frac{E}{G_2} \quad \text{amperes} \tag{18-30}$$

Values of G for single-conductor and G_2 for 3-conductor cable may be taken from Table 18-22.

Geometric Factors of Cables. See Table 18-22. Intermediate values may be found by interpolation.

TABLE 18-22 Table of Geometric Factors of Cables

Ratio $\frac{T+t}{d}$	G Single conductor	Sector factor	Three-conductor cables: G_1 at ratio t/T 0	0.5	1.0	G_2 at ratio t/T 0	0.5	1.0
0.2	0.34	. . .	0.85	0.85	0.85	1.2	1.28	1.4
0.3	0.47	0.690	1.07	1.075	1.08	1.5	1.65	1.85
0.4	0.59	0.770	1.24	1.27	1.29	1.85	2.00	2.25
0.5	0.69	0.815	1.39	1.43	1.46	2.10	2.30	2.60
0.6	0.79	0.845	1.51	1.57	1.61	2.32	2.55	2.95
0.7	0.88	0.865	1.62	1.69	1.74	2.55	2.80	3.20
0.8	0.96	0.880	1.72	1.80	1.86	2.75	3.05	3.45
0.9	1.03	0.895	1.80	1.89	1.97	2.96	3.25	3.70
1.0	1.10	0.905	1.88	1.98	2.07	3.13	3.44	3.87
1.1	1.16	0.915	1.95	2.06	2.15	3.30	3.60	4.05
1.2	1.22	0.921	2.02	2.13	2.23	3.45	3.80	4.25
1.3	1.28	0.928	2.08	2.19	2.29	3.60	3.95	4.40
1.4	1.33	0.935	2.14	2.26	2.36	3.75	4.10	4.60
1.5	1.39	0.938	2.20	2.32	2.43	3.90	4.25	4.75
1.6	1.44	0.941	2.26	2.38	2.49	4.05	4.40	4.90
1.7	1.48	0.944	2.30	2.43	2.55	4.17	4.52	5.05
1.8	1.52	0.946	2.35	2.49	2.61	4.29	4.65	5.17
1.9	1.57	0.949	2.40	2.54	2.67	4.40	4.76	5.30
2.0	1.61	0.952	2.45	2.59	2.72	4.53	4.88	5.42

Example. Find 60-Hz charging kVA for 33-kV cable having three 350,000-cmil sector-type conductors each with 10/32 in of paper and a 5/32-in belt.

$$T = 0.313 \text{ in} \qquad t = 0.156 \text{ in} \qquad d = 0.681 \text{ in} \qquad t/T = 0.5$$

$$(T + t)/d = (0.313 + 0.156)/0.681 = 0.69; \; E = 33/1.73 = 19 \text{ kV}$$

Interpolating in Table 18-22, we find $G_2 = 2.78$.

For sector-type cable, G_2 must be multiplied by the sector factor for 0.69, which is seen to be 0.86 in the sector-factor column in Table 18-22. For such a cable,

$$G_2 = 0.86 \times 2.78 = 2.39 \qquad \text{and} \qquad I = (0.063 \times 19)/2.39 = 0.5\text{A}/1000 \text{ ft}$$

Charging kVA $= 3IE = 3 \times 0.5 \times 19 = 28.5$ kVA/1000 ft, and for a cable having a length of 20 mi it would be $20 \times 5.28 \times 28.5 = 3010$ kVA. For single-conductor cables, $t = 0$ and $(T + t)/d = T/d$, which is used to get the value of G from the values for single-conductor cable in Table 18-22.

Cable Terminations. A cable termination must perform several functions:

1. Provide means for electrical connection of the cable to an equipment or circuit.
2. Control the electrostatic stresses so that there is no electrical-discharge activity in the termination at design voltage levels. One important consideration is to control the voltage stresses where the change is from a uniform radial field within the (shielded) cable to a new configuration beyond the termination of the shield. Other considerations are to provide adequate flashover and creepage strength to nearby grounds.
3. Prevent loss of gas or liquid insulation impregnant from the cable, where needed, or from the termination.

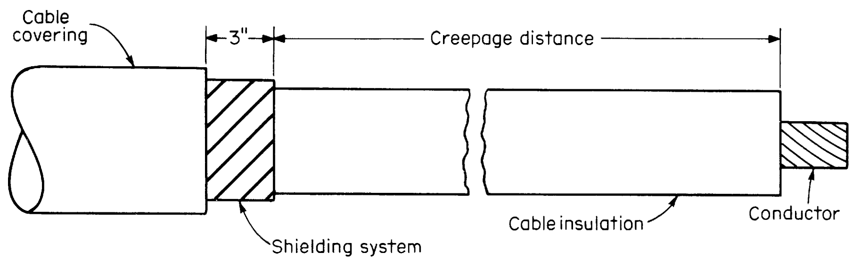


FIGURE 18-56 Elementary plain-shield termination.

4. Provide suitable mechanical and/or hermetic termination of the sheath, where used.
5. Serve as a load-break switch or separable connection, where needed.

Many types of terminations are in use, ranging from those made by hand in the field to factory-made types requiring very little work in the field. Two broad classifications are live-front and dead-front. The former involves exposed, bare electrical connections and possibly lengths of unshielded cable. The latter type of termination is completely enclosed in a semiconducting or metallic structure essentially at ground potential, such that it can be touched without hazardous shock while the equipment is energized.

Elementary Stress-Cone Termination. Figure 18-56 shows a single-conductor shielded cable prepared for termination. Figure 18-57 shows an elementary stress-cone termination. The stress cone may be built up by using tapes of a material compatible with the cable insulation, or it may be a factory-molded stress cone which is slipped over the (solid-dielectric) cable insulation system after the end of the cable has been properly dressed.

The stress cone serves to keep dielectric stresses at acceptable values. Without the stress cone, the plain termination of Fig. 18-56 would experience excessive stresses at normal voltage near the end of the shielding system, leading quickly to failure. Occasionally, an overall cover tape may be provided. Usually, a compression-type connector is installed on the cable conductor to provide a means of electrical connection to the equipment.

Separable Insulated Connectors. Accompanying the rapid growth in the use of single-conductor, concentric-neutral (or shielded cables employing solid-dielectric insulation) has been the development of separable insulated connectors of the dead-front classification. A typical termination consists of

1. An elbow, as shown in Fig. 18-45, which is physically and electrically connected to the end of a properly dressed cable.
2. A load-break switch module where the load-break function is specified (see Fig. 18-44).
3. An apparatus bushing. When a load-break module is not used, the elbow is mated directly to the apparatus bushings.

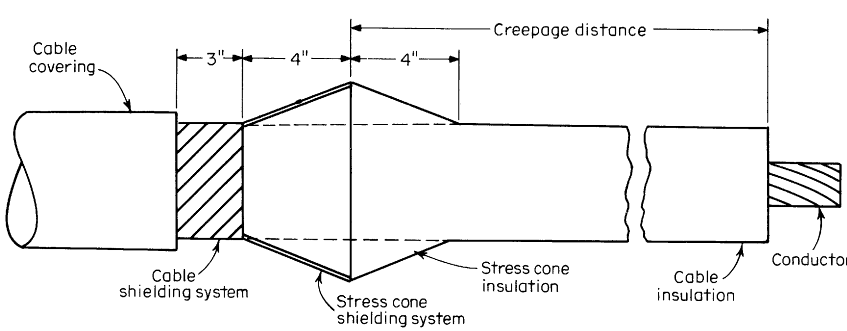


FIGURE 18-57 Elementary stress-cone termination.

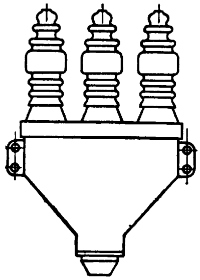


FIGURE 18-58 Three-conductor disconnecting-type pothead.

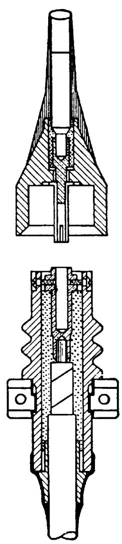


FIGURE 18-59 Disconnecting-type pothead.

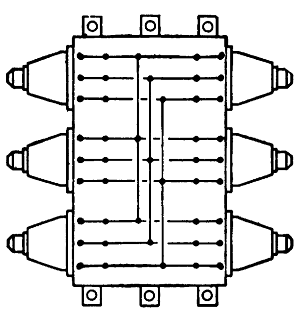


FIGURE 18-60 Six-way disconnecting cable-junction box.

The external shield and the insulation of the elbow and switch module are made of synthetic rubber. The stress-relief function is designed into the molded configuration.

When the elbow is in its normal connected position, the termination is dead-front and submersible.

A wide variety of accessory components are available, including multitaps, insulated bushings, feed-through bushings, insulating caps, etc. so that wide flexibility in operating procedures can be obtained for differing system and equipment configurations.

Separable insulated connectors are available in 200-A ratings for use on grounded-wye systems of the 15-, 25-, and 35-kV classes. Ratings of 600 A also are available, but these usually cannot be opened under load or while the cable is energized.

To install the elbow, the end of the cable is dressed according to the manufacturer's specifications; often a jig is used so that proper removal of semiconducting shield, correct dressing dimensions, exposure of the conductor, etc., are easily obtained. A crimped connector is then installed on the cable conductor, and the elbow is then slipped over the cable, internal male contact installed, and the cable or concentric neutral is connected to the semiconducting outer shield of the elbow.

Obviously, a termination of this kind can be installed much more quickly and with a lower level of skill needed than in terminating the traditional paper-insulated, lead-sheathed cables.

Pothead Terminations. Where cables are connected to overhead systems or to switchgear equipment, they often are terminated by means of potheads, such as that shown in Fig. 18-58. An extremely wide variety of types is in use, depending on

1. System voltage.

2. Type of cable insulation.

3. Single- or multiconductor cable.

4. Type of jacket or sheath.

5. Whether "live" side of pothead is outdoor or within an equipment. Some potheads are of the disconnecting type as indicated in Fig. 18-59.

There is an increasing trend toward the use of factory-made molded-rubber terminations in place of the traditional potheads, particularly with cables having solid-dielectric insulation.

Subway Junction Boxes. Subway junction boxes are sometimes used in cable systems to interconnect distribution circuits at points where it is desired that the connection be opened, at times, for construction or operating purposes and where overhead disconnecting facilities are not available. In low-voltage systems, such as 208Y/120-V secondary networks, such boxes may include sectionalizing fuses or connecting links. They are used in a number of different circuit configurations, a six-way junction box being shown in Fig. 18-60.

Splicing. Cable splices to a great extent resemble "back-to-back" portions of cable terminations. The completed splice provides

1. Electrical connection between the cable conductors, usually by means of crimped connectors.
2. Insulation over the exposed conductors and connector.
3. Jointing of the shielding or concentric neutral systems of the two cable sections so that electrical stresses are properly controlled.
4. Jointing of the jacketing systems or sheaths.
5. A "stop" function where the two cable insulation systems are different, for example, oil-impregnated paper on one side and solid dielectric on the other. Here it is necessary to contain the oil in the paper insulation and to exclude its penetration into the solid dielectric.
6. A disconnecting function, where required.

Factory-made splices are used extensively for splicing cables with solid-dielectric insulation. Where the disconnecting function is required, various combinations of multitaps and elbow terminations are used.

The traditional method of splicing paper-insulated, lead-sheathed cable is shown in Fig. 18-61. This method, which employs hand-wrapped insulating tapes, requires a high level of skill and training as contrasted to the use of factory-made splices for joining solid-dielectric cables. In jointing single-conductor cables, the lead sheath is removed about 6 in back from the end, and enough insulation is cut away to permit a soldered connection to be made. When the connection is complete, the bare parts are wrapped with tape until the equivalent of cable insulation has been applied. A lead sleeve which has previously been slipped over one of the cables is now wiped on the two cable sheaths so as to enclose the joint. Air space around the joint is then filled with hot insulating compound poured into a small hole in one end of the sleeve; a similar hole is left in the other end to allow air to escape. These holes are then closed by soldering. The joint should be allowed to cool before it is moved, so that the compound will hold the parts rigidly in place.

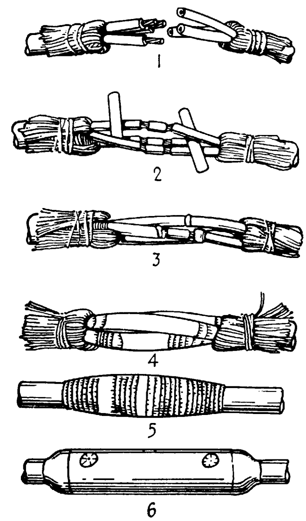


FIGURE 18-61 Successive stages in cable splicing.

In jointing 3-conductor cables, the lead must be removed about 10 in to facilitate taping the conductors (see Fig. 18-61). In making joints for 6600 V and higher, it is important that as little air remain in taping as possible. If paper tape is used, each layer should have compound poured over it before the next is applied.

Installation of Cable. Generally, *direct-buried cable* in underground residential distribution (URD) systems is buried in a trench, usually 36 in or more deep. Often, random lay of the power and telephone cables is employed, with no intentional separation. When soil conditions permit, the trench is backfilled with the original material. In some cases a selected backfill and/or protective covering over the cables may be necessary.

Where URD circuits must be routed under streets or other paved surfaces, or in locations such that it would be impractical to dig in order to repair a faulted cable, duct installation often is used.

Where soil conditions and circuit configurations are favorable, it is possible to directly plow in the cable by means of a special plow which breaks the earth ahead of the cables and guides them into the furrow.

In *duct installations*, the most common method of preparation has been to use wood or metal rods which are pushed into the duct section by section as they are connected together. When the opposite end of the duct is reached, a wire is attached to the end of the last rod and is pulled into the duct as the rods are withdrawn.

When the duct is airtight, a piston with attached flexible wire can be blown through the duct by means of compressed air. This method is quicker and less laborious than the rodding process.

Cables are pulled through the duct by means of a pulling line, usually wire rope, and a power-driven cable-pulling winch. Cables of moderate size and length usually are pulled by means of cable grip, which is a type of woven-wire basket designed to increase its grip on the cable as the tension increases. Often a flexible pull-in tube is used in the manhole to prevent damage to the cable being pulled and to other exposed cables in the manhole.

With long sections or larger cables, it may be necessary to use a pulling eye rather than a cable grip because the pulling tension may exceed the capability of the cable grip. The pulling eye is a steel eye which usually is fastened directly to the cable conductors.

When a new cable is to be installed in an existing duct, it is generally desirable that the diameter of the duct be at least 3/4 in greater than that of the cable. Where the duct section is exceptionally long or contains relatively sharp bends, a clearance of 1 in may be needed. On short, straight sections, 1/2-in clearance may be acceptable. Where several single-conductor cables are to be drawn into the same duct, the cable reels are set up in tandem and all cables are pulled into the duct simultaneously.

Cable Training. The location of cables in manholes is determined primarily by the ducts which they occupy. The cables should be fanned out as they leave their ducts so that they do not cross other cables or ducts. Sufficient length must be left in manholes to permit training on racks around the manhole walls, as shown in Figs. 18-52 and 18-53, and for joining. Radius of bends should be greater than the minimum safe bending radius for the cable, and cable movement also must be considered. The safe bending radius varies with the size, type of sheath or armor, etc. and generally is on the order of 8 to 12 times the overall diameter of the cable for power cables.

With large cables, the periodic load cycles cause repeated flexing and movement in the manholes, and duct-mouth protection is often used to prevent cracking of lead sheaths. This may consist of a piece of galvanized metal inserted under the cable and arranged to prevent the sheath from being pressed against the sharp edges of the duct mouth.

In order to limit damage resulting from cable faults to the failed cable, it is quite common to fireproof the cables in a vault or manhole. Such fireproofing usually is done with asbestos tapes and asbestos or mortar cements.

Sheath Bonding. With cables having a lead or other metallic sheath, it is common practice to bond together the different sheaths in a manhole or vault. Bonding consists of electrically connecting together the various sheaths. Various types of bonding systems are in use to maintain the sheaths at a common potential near ground potential, thus reducing the danger to workers who may be in the manhole when a cable fault occurs. This also eliminates the possibility of serious arcing occurring between the sheaths of the faulted and unfaulted cables.

Selection of Duct Position for Cable. Cables used in local distribution should be installed in the top row of ducts so that manholes for service connections and lateral circuits can be of a relatively shallow construction. In distribution manholes, the higher-voltage cables and cables of through lines should be placed in the lower and outside ducts, where possible, and they should be trained with the least possible interlacing with other cables.

Ambient Earth Temperatures. Ambient earth temperatures vary with geographic location, season, and depth. Average daily air temperature at a given location follows a more or less sinusoidal curve over the seasons of the year. As a result, the earth ambient temperatures also exhibit an annual sinusoidal variation. At depths greater than 1 to 2 ft, there is essentially no daily variation in earth ambient, but there is a seasonal variation which is greater at shallow depths, decreasing with depth. In addition, as the depth increases, the variation in ambient temperature increasingly lags behind the daily ambient air temperature curve; at a depth of 6 ft, this lag may be as great as 6 to 8 weeks. At depths of 20 to 30 ft, the earth temperature remains practically constant at about the mean annual air temperature. As a result, the earth temperature tends to increase with depth in the winter and decrease with depth in the summer. At a depth of 3 1/2 ft, typical temperatures are as follows:

	Temperature, °C	
	Summer	Winter
Northern U.S.	20–25	2–15
Southern U.S	25–30	10–20

Calculation of Ampacities of Cables. The precise calculation of ampacities of cables is extremely complex and has been the subject of numerous technical papers. This complexity is due not only to the characteristics of the thermal circuit, such as heat transfer through the cable insulation and sheath, transfer to duct or earth, and transfer from duct bank to earth but also to the fact that losses in the cable conductor are subject to skin and proximity effects. Also, additional losses can occur in the cable-shielding system depending on the nature of the installation.

Presently accepted methods of calculation, empirical data, and numerous references are treated in the following references:

1. IEEE Standard 835-1994, *IEEE Standard Power Cable Ampacity Tables*, goes into great detail explaining where the numbers came from. The types of cables for which ampacities have been calculated range from simple crosslinked polyethylene insulated 600-volt cables up to various medium voltage cables. The technical introduction covers cable construction, installation conditions, calculation methods, calculation examples and details of the assumptions made in calculating the ampacities that appear in the tables.

2. Ampacities, Including the Effect of Shield Losses for Single Conductor Solid Dielectric Power Cable 15 kV through 69 kV, NEMA WC 50-1976/ICEA P-53-426, 2nd ed. (R1993, R1999).

Maximum Allowable Conductor Temperature. The ICEA temperature ratings for polyethylene (thermoplastic) and cross-linked-polyethylene-insulated power cables are

	Max. conductor temperature, °C	
Insulation	Normal operation	Emergency overload
Polyethylene	75	90
Cross-linked polyethylene	90	130

Maximum conductor temperatures for impregnated paper-insulated cables are given in Table 18-23, as adapted from Publication P-46-426 of the IPCEA.

Ampacity of Cables. There is a growing use of single-conductor, solid-dielectric power cables for important 3-phase distribution circuits in the 15- to 35-kV class. Various shielding systems are in use including concentric wires, ribbons, and tapes. In many cases, on 4-wire primary-distribution circuits, the concentric wires are used as neutral conductors. When the shields are bonded together and grounded at multiple locations, circulating currents can flow in the shields, resulting in I^2R losses and appreciable heating effect. Such losses may be significant when the cables are spaced.

The AIEE-IPCEA ampacity tables do not include the effects of circulating-current losses, but these effects are included in the ampacity tables of IPCEA Publication P-53-426, NEMA Publication WC50-1976, "Ampacities Including Effect of Shield Losses for Single-Conductor Solid-Dielectric Power Cable 15 kV through 35 kV (Copper and Aluminum Conductors)." The ampacity data in Table 18-24 have been taken from this publication and apply to directly buried, solid-dielectric power cable

Table 18-25 has been adapted from the IPCEA Publication P-53-426 (and NEMA Publication WC50-1976) to illustrate typical ampacities of single-conductor, solid-dielectric power cables installed in underground ducts. The type of installation is assumed to be directly buried fiber or plastic duct of inside diameter nominal pipe size to provide a minimum diametral clearance of 0.75 in between cable outside diameter and inside diameter of the duct. The assumed arrangement of the ducts is shown in Fig. 18-62.

TABLE 18-23 Maximum Conductor Temperatures for Impregnated-Paper-Insulated Cable

	Conductor temperature, °C		
Rated voltage, kV	Normal operation	Emergency operation	
Solid-type multiple conductor belted			
1	85	105	
2–9	80	100	
10–15	75	95	
Solid-type multiple conductor shielded and single conductor			
1–9	85	105	
10–17	80	100	
18–29	75	95	
30–39	70	90	
40–49	65	85	
50–59	60	75	
60–69	55	70	
Low-pressure gas-filled			
8–17	80	100	
18–29	75	95	
30–39	70	90	
40–46	65	85	
Low-pressure oil-filled and high-pressure pipe type			
		100 h	300 h
15–17	85	105	100
18–39	80	100	95
40–162	75	95	90
163–230	70	90	85

Source: Copyright 1962 by Insulated Power Cable Engineers Association. Used by permission.

18.25 FEEDERS FOR RURAL SERVICE

Basic Conditions. Rural distribution differs from urban in that consumers are farther apart and load units are generally small. Since distances are great, the primary system voltage should be the 15-kV class or higher, and the load per mile being low requires the cost of feeder construction to be as low as is consistent with a reasonable degree of permanence and reliability. One transformer per customer is required in many cases. Rural construction since the late 1930s has made electric service available to practically every farm. Efforts are directed now to bolstering capacity to serve the growing loads.

Poles and Spans. Design of overhead lines for rural service differs from that of urban lines in several respects. Costs are reduced by using longer spans and as few accessories as possible. Longer spans mean greater sag and higher poles to get proper clearance at the low point of the span. The increase in sag may, however, be reduced by use of higher tensile stresses in conductors. This is possible when steel is employed in conjunction with copper or aluminum wires. Steel is combined with copper in a high-strength strand known as Copperweld, which has 30% or 40% of the conductivity of a copper conductor of equal size, or in a similar aluminum and steel conductor known as Alumoweld. When greater conductivity is needed, one or more strands of hard copper are stranded with or around the Copperweld or hard aluminum strands around Alumoweld. Steel is also stranded

TABLE 18-24 Ampacity of Single-Conductor Solid Dielectric Power Cable Installed Direct Buried

Cond. size	Neutral size	Ampacity	°C*	W/ft^2
1/0	Full	241	66	46.2
1/0	1/2	245	66	45.2
1/0	1/3	246	66	44.8
1/0	1/6	247	66	44.2
4/0	Full	339	71	47.4
4/0	1/2	349	70	45.7
4/0	1/3	355	69	44.8
4/0	1/6	361	69	43.6
500	1/3	513	75	45.5
500	1/6	544	74	43.1
500	1/12	566	73	40.9
500	1/18	575	72	40.2
750	1/3	575	75	42.3
750	1/6	624	75	40.8
750	1/12	671	74	38.5
750	1/18	690	73	37.2
1000	1/6	675	76	39.8
1000	1/12	748	76	37.9
1000	1/24	799	75	35.6
1000	1/36	819	74	34.8

90°C aluminum conductor.
Single circuit—three cables spaced.
25°C earth ambient.
90 rho soil resistivity.
*Corresponding earth interface temperature in °C.
Source: From IEEE Std 835-1994 (© 1994 IEEE).

with aluminum wires into ACSR conductor. Such types of conductor have ample conductivity for rural lines, and they have been used widely. In level country, spans of 400 to 600 ft are practical, while in hilly country, spans of 800 to 900 ft are occasionally possible.

Cable. Design of underground circuits for rural service is similar to that of urban underground circuits. Concentric-neutral cables are likely to be used in both types of systems. Because the rural circuits have longer uninterrupted runs of cable, there is a better opportunity to plow in the cable rather than digging trenches. Plowing results in lower installed costs per unit length of cable. In fact, some electric suppliers report that the total cost of a rural underground system is less than the cost of an overhead system to serve the same load.

Location of Circuits. Rural-service circuits are run along main highways, where the largest number of users may be reached. Branches along intersecting roads are extended as may be warranted by service requirements. In some cases, private rights of way, maintained for transmission lines, may be utilized.

Voltage. Rural circuits may be extended 5 to 50 mi from the point of supply, and voltage used for primary distribution must be chosen accordingly. Loadings are often so small that the minimum size of conductor required for dependable strength for overhead or cable insulation for underground is sufficient to meet requirements of voltage drop and line loss. This is particularly true when the higher voltages are used.

The most common primary voltage used in rural areas is 12,470Y/7200 V, 4-wire for normal load densities and 24,940Y/14,400 V for very light load densities. There is a trend toward using both 24,940Y/14,400 V and 34,500Y/19,900 V for all types of rural areas.

TABLE 18-25 Ampacity of Single-Conductor Solid Dielectric Power Cable Installed in Underground Ducts

Cond. size	Neutral size	Ampacity 75% LF	Ampacity 100% LF
250	1/3	272	239
250	1/6	279	245
250	1/12	282	249
250	1/18	284	250
350	1/3	316	277
350	1/6	329	288
350	1/12	337	295
350	1/18	339	298
500	1/3	364	317
500	1/6	386	337
500	1/12	402	351
500	1/18	407	356
750	1/3	414	359
750	1/6	449	389
750	1/12	480	417
750	1/18	493	428
1000	1/6	490	424
1000	1/12	536	464
1000	1/24	571	494
1000	1/36	584	506

90°C aluminum conductor.
Two circuits—three cables spaced 7.5 inches.
25°C earth ambient.
90 rho soil resistivity.
Source: From IEEE Std 835-1994 (© 1994 IEEE).

Single-phase circuits are most economical for the usual light loads found in rural areas and where power units do not exceed 10 hp. Vee-phase circuits consisting of 2-phase conductors and the neutral are an economical method of supplying 3-phase loads using open-wye-open-delta transformer banks. Full 3-phase, 4-wire construction will be desirable for many areas. In some cases there may be relatively small 3-phase loads in a single-phase area. Often these loads can be supplied economically from a single-phase system by means of a phase converter, the output of which is 3-phase voltage.

Limitations of voltage and distance are illustrated by the following table showing kilowatt-miles corresponding to a 5% line drop at 80% power factor for a circuit of 1/0 ACSR, or its equivalent in other metals.

Kilowatts × miles for 5% voltage drop, power factor 80%

System	4.16 kV	12.47 kV	24.94 kV	34.5 kV
Single-phase	82	737	2949	5646
3-phase	488	4375	17919	33815

Values for other sizes are approximately in proportion to relative cross section. In order to determine specific voltage-drop values for 3-phase overhead and underground circuits, refer to Table 18-3.

Conductors and Spans. Because of the economy of using long spans, the choice of span lengths and conductor strength is of much importance in planning rural lines. Single-phase lines are commonly taken from a 3-phase system with neutral grounded. The grounded conductor is carried on a bracket about 2 ft below the phase wire, which rests on an insulator carried on the top of the pole. No crossarm is required, except on a main line of more than one phase.

While the conductivity of No. 4 ACSR or Copperweld may be thermally adequate for the greater part of a rural system, system economics can dictate the use of larger conductors. The strength of No. 4 ACSR or Copperweld is usually ample for spans of 350 to 600 ft, depending upon design-loading conditions. Conductors should be sagged in accordance with the conductor manufacturer's recommendations.

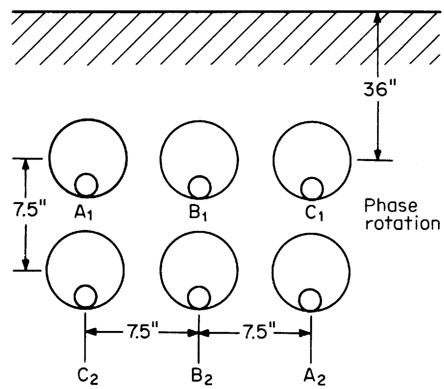


FIGURE 18-62 Arrangement of ducts.

Poles. The strength of poles should be determined for the height required by the methods described in Secs. 18.14 and 18.15. The length of poles required in any situation must be such as to allow for depth of setting and height of wire supports needed to give proper clearance above ground at the low point of the span. Such clearances should be not less than value shown in the NESC. In the case of road and railroad crossings, the necessary clearance may sometimes be more readily had by placing one end of the span near the crossing, thus avoiding having the low part of the span over the crossing. In rolling or hilly areas, it is desirable to locate poles on higher elevations to permit use of longer spans and greater sags.

Where no ice loading is likely, 30-ft poles can be used for two conductors of a single-phase branch on level ground or on long even slopes with span lengths to 400 ft. This often will preclude, however, the addition of communication circuits to the poles. Where ice and wind loading is expected with some regularity, it is necessary to use 35-ft poles for spans exceeding 300 ft. At corners or angles, poles should be supported by guying or bracing to support unbalanced longitudinal stresses.

Crossarms or equivalent equipment are required for the main 3-phase circuits and for lines supplying any user taking 3-phase service for power. A two-pin arm is often used with the third phase on the pole top. The grounded neutral is carried on the side of the pole about 2 ft below the arm.

Transformer Installations. Transformers usually supply not more than one or two customers, and sizes, therefore, are small compared with the average used in urban work, 10 to 15 kVA being average for single-phase installations. Where points of use are more than about 500 ft apart, it is usually most economical to provide separate transformers. When two users are within this distance, they can be served by placing a transformer between them and constructing a secondary. Rural loads on some systems have grown to the point where 15- and 25-kVA transformers are required.

Transformer capacity usually may be selected on the basis of loading the transformer to 150% of nameplate rating for peak loads lasting for 1 to 2 h. Pumping for drainage or irrigation is likely to require rated capacity more nearly equal to load.

Stray Voltages. Stray voltages on dairy farms may cause lowered milk production and increased mastitis in dairy cattle. Dairy cattle are particularly sensitive to low magnitudes of voltage. Voltages on the order of 0.5 V occurring between metal stanchions or metal drinking cups and the concrete floor may be troublesome. It should be noted that the same symptoms may be due to other causes and that stray voltages are not always to blame.

One characteristic of the common-neutral distribution system, in which the neutral conductor is common to both the primary and secondary systems, is the multiplicity of ground connections between the neutral conductor and earth. Unbalanced load conditions on either the farm secondary system or the utility primary system result in current flow in the neutral conductor. Due to the multiplicity of ground connections, some portion of the neutral current flows in earth. These earth currents cause stray voltages to appear in the earth. Deteriorated insulation on farm wiring and machinery can also cause earth currents, and these causes should be eliminated first.

One way to minimize stray voltages in the dairy barn is to cast wire mesh into the concrete floor and to bond the mesh and all metal structures together to establish equal potentials. In cases where current flow in the utility primary neutral is identified as a cause of stray voltages, it may be necessary to isolate the primary and secondary neutrals of the distribution transformer serving the farm. The NESC, IEEEC2-1997, addresses this situation in Section 97D.

18.26 DEMAND AND DIVERSITY FACTORS

Demand Factor. The ratio of maximum demand to total load connected, expressed as a percentage, is termed the *demand factor* of an installation. For example, if a residence having equipment connected with a total rating of 6000 W has a maximum demand of 3300 W, it has a demand factor of 55%. Demand factors of various types of large loads are helpful in designing systems, particularly those in buildings. As an example, a single household electric clothes dryer, of course, has a demand factor of 100%, but 25 dryers in a group have a demand factor of 33%. Similarly, three to five all-electric apartments in a multifamily dwelling have a demand factor of 45%. The lower the demand factor, the less system capacity required to serve the connected load. However, summer air conditioning and winter electric heating are loads that make for high demand factors.

Coincidence or Diversity Factor. The *coincidence factor* is defined as the ratio of the maximum demand of the load as a whole, measured at its supply point, to the sum of the maximum demands of the component parts of a load. The *diversity factor* is the reciprocal of the coincidence factor. Coincidence factors can be applied to known consumer demands for estimating the loading of distribution transformers, lines, and other facilities. Coincidence factors for residential consumers can vary over a wide range for different types of consumers. The coincidence factor for a large group of consumers with no major appliance might be as low as 30%, whereas a group of electric-heating consumers might be as high as 90%.

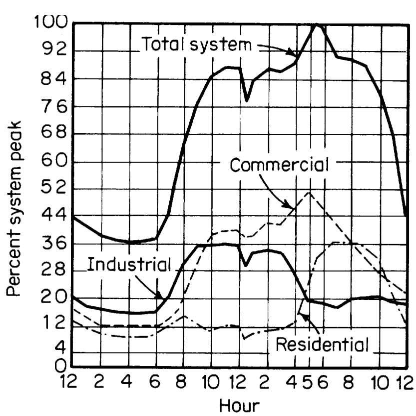


FIGURE 18-63 Characteristic metropolitan load pattern.

Diversity Between Classes of Users. The *daily-load curve* of a utility is a composite of demands made by various classes of users. The load curve on the day of maximum total system peak occurs when class loads gang up to create this maximum demand for the year. This is not necessarily the day, and usually is not the day, of any particular class peak. Class load curves on the day of system peak are illustrated in Fig. 18-63.

Air-conditioning loads have shifted these curves for many systems to cause daytime peaks during hot weather in the summer. Electric house heating builds heavy morning and evening loads during cold weather in the winter.

Diversity in the Feeder System. The diversity of demands of transformers on a radial feeder makes the maximum load on the feeder less than the sum of the transformer loads. The diversity factors of a feeder vary greatly depending on load conditions. Some typical diversity factors are given in Table 18-26. The diversity factor of lighting feeders ranges from 1.1 to 1.5, while that of

TABLE 18-26 Diversity Factors

Elements of system between which diversity factors are stated	Diversity factors for			
	Residence lighting	Commercial lighting	General power	Large users
Between individual users	2.0	1.46	1.45	
Between transformers	1.3	1.3	1.35	1.05
Between feeders	1.15	1.15	1.15	1.05
Between substations	1.1	1.10	1.1	1.1
From users to transformer	2.0	1.46	1.44	
From users to feeder	2.6	1.90	1.95	1.15
From users to substation	3.0	2.18	2.24	1.32
From users to generating station	3.29	2.40	2.46	1.45

mixed light-and-power feeders is likely to be 1.5 to 2 or more. At the substation there is also a diversity factor of 1.05 to 1.25 between the sum of feeder maxima and the substation maximum. A large system has a further diversity factor between substations of 1.05 to 1.25. Total diversity factors in a large system are somewhat as in Table 18-26.

18.27 DISTRIBUTION ECONOMICS

Economic Comparisons. The most straightforward and generally applicable technique to use in distribution system investment problems is that of making economic comparisons on the basis of the present value of all future annual costs. That is, the economic choice is the one with the lowest present value of all future costs. With this as a criterion, the procedure for making an economic comparison between alternatives is a simple two-step operation, that is,

1. Estimate for each alternative the annual costs for each year.

2. If annual costs are not uniform, calculate their present value.

Time Value of Money. Money does have time value, and rent or interest on its use has to be paid. It is obvious that an alternative which requires the least expenditure immediately would be best, everything else being equal.

The process of taking money and finding its equivalent value at some future time is called a *future worth* or *future value* calculation. This calculation is the same as that used in determining the effect of compound interest.

If 8% is the established interest rate, then \$100 today is equivalent to \$100(1 + 0.08) or \$108 a year from now, and $100(1 + 0.08) + 100(1 + 0.08) \times 0.08 = 100(1 + 0.08)^2$ 2 years from now and $100(1 + 0.08)^{10}$ 10 years from now. The expression $(1 + i)^n$ is called the *single-payment compound amount factor*, where i is the interest rate and n is the number of years. These factors and others are readily available for various interest rates and number of years in economic books such as *Principles of Engineering Economy* by Eugene L. Grant.

It should be noted that the use of 8% for the interest rate is for illustrative purposes only. The actual interest rate to use will be determined by the economic conditions at the time.

Hence, to find the future worth of \$100, 10 years later in the preceding example, first look up the compound amount factor in the 8% interest table for year 10; then multiply it by 100. The compound amount factor for this case is 2.159, and the future worth calculates to be 100(2.159) = \$215.90.

The process of finding the equivalent value of money at some earlier time is called a *present worth* or *present value* operation. The present worth calculation is the reverse of the future worth

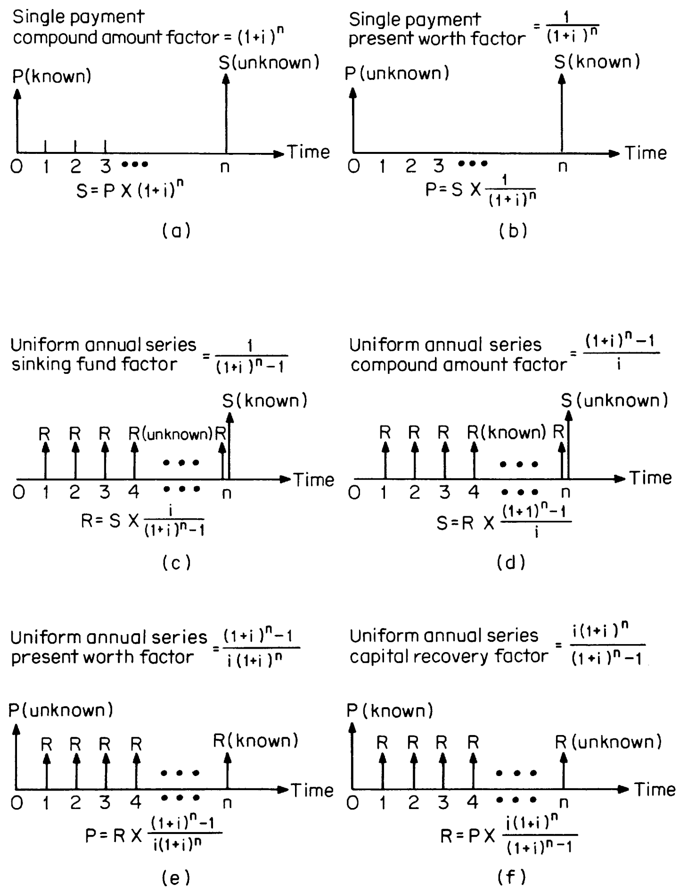


FIGURE 18-64 Graphic interpretation of compound interest factors.

calculation. If \$100 today has a future worth a year from now of \$108, then we can also say that \$108 a year from now has a present worth of \$100 today. The present worth factor is the inverse of the future worth factor, and it also may be found in interest tables. Since the future worth factor for n years is $(1 + i)^n$, where i is the interest rate, the present worth factor is $1/(1 + i)^n$.

To determine the present worth, as of today, of a \$100 cost anticipated to be incurred 2 years from now where the interest rate is 8%, first the present worth factor of 0.8573 is obtained from interest tables. Then multiplying this factor by \$100 gives the present worth of \$100(0.8573) = \$85.73.

Formulas for calculating the compound interest factors and a graphic interpretation of these factors are shown in Fig. 18-64.

Annual Charges. It is desirable to have a convenient method of calculating the annual costs of capital investments made in an alternative scheme. Fortunately, this can be done by using a level carrying charge which is expressed as a percentage of the original investment.

The total revenue requirements of a piece of equipment are the sum of the annual charges for

1. Return on investment
2. Depreciation
3. Income tax

4. Property taxes
5. Insurance
6. Operating and maintenance expenses

The first five of these charges can be conveniently estimated as a percentage of original investment. The operating and maintenance charges should be estimated separately for each project because they do not relate to capital investment as a percentage.

Level Annual Carrying Charges. The level annual carrying charge is the percentage by which the capital investment can be multiplied to determine its annual cost on a uniform basis. The value of this carrying charge is very much dependent on the expected life of the piece of equipment because depreciation varies in accordance with life expectancy. A method of obtaining the level annual carrying charge is as follows: (1) calculate the sum of the annual charges for return on investment, depreciation, income tax, property tax, and insurance for each year of the expected life of the piece of equipment; (2) use the appropriate present worth factor with each annual cost to convert the annual cost to a present worth value; (3) sum up these values to obtain the total present worth of the annual carrying charges; and (4) multiply the total present worth by the capital recovery factor (see Fig. 18-64) to get the equivalent uniform annual charge. Figure 18-65 shows graphically the actual and equivalent carrying charges for a capital investment of a piece of equipment with a 5-year life and an assumed 8% cost of money.

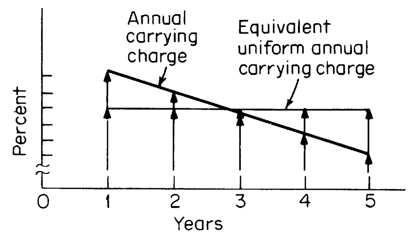


FIGURE 18-65 Representation of carrying charges.

The total carrying charges with 8% cost of money for various service lives are estimated as follows:

Years of life	Level annual total carrying charge in %
5	30.82
10	20.59
15	17.44
20	16.04
25	15.34
30	14.96
35	14.76
40	14.67
45	14.63
50	14.63

Operating and Maintenance Expenses. This cost component varies with the nature of the project. It is usually not a direct function of the capital invested and may have an inverse tendency. That is, alternatives often exist for higher capital expenditures to reduce operating costs. Therefore, it is *not* expressed as a percent of capital investment in most cases. Nevertheless, it should be included in annual costs.

Study Period. When determining the economic comparison of alternates by comparing the present worth of annual costs, the study period should be taken to the point that the alternates are equivalent in capability. If this is not practical, the study should be taken so far into the future that the difference in present worth would be insignificant.

Economic Evaluations. A simple example will show a comparison between two alternatives. Let *CC* represent the capital investment multiplied by the level annual carrying charge, *O&M* represent

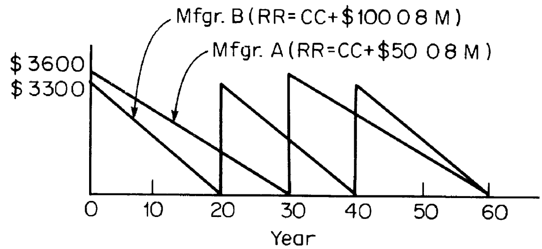


FIGURE 18-66 Time diagram.

annual operation and maintenance, and *RR* represent the total revenue requirement necessary annually to carry the project. A pad-mounted sectionalizing switch is needed for an underground circuit. The choice is between two manufacturers who can supply the switch but with different characteristics as follows:

	Mfr. *A*	Mfr. *B*
Installed cost	$3600	$3300
Operating and maintenance	50/year	100/year
Expected life	30 years	20 years

There is no salvage value at end of life. Determine which alternative is less expensive.
The first step is to draw a time diagram like Fig. 18-66. The common point in time for the two alternatives is 60 years, so two cycles of *A* should be compared with three cycles of *B*.

Present worth analysis:

$$\begin{aligned}\text{PW mfr. } A \text{ alternative} &= 3600 \times 0.1496 \times 11.258 + 50 \times 11.258 \\ &\quad +(3600 \times 0.1496 \times 11.258 + 50 \times 11.258)\, 0.0994 \\ &= 6063.11 + 562.90 + 658.63 \\ &= 7284.64\end{aligned}$$

$$\begin{aligned}\text{PW mfr. } B \text{ alternative} &= 3300 \times 0.1604 \times 9.818 + 100 \times 9.818 \\ &\quad +(3300 \times 0.1604 \times 9.818 + 100 \times 9.818)\, 0.2145 \\ &\quad +(3300 \times 0.1604 \times 9.818 + 100 \times 9.818)\, 0.0460 \\ &= 5196.86 + 981.80 + 1325.32 + 284.22 \\ &= 7788.20\end{aligned}$$

where 3600 = installed cost of mfr. *A* switch
0.1496 = level annual carrying charge for 30-year *A* switch
11.258 = 8%, 30-year uniform annual series present worth factor
50 = *O&M* of *A* switch
0.0994 = 8%, 30-year single-payment present worth factor
3300 = i.nstalled cost of mfr. *B* switch
0.1604 = level annual carrying charge for 20-year *B* switch
9.818 = 8%, 20-year uniform annual series present worth factor
100 = *O&M* of *B* switch
0.2145 = 8%, 20-year-single payment present worth factor
0.0460 = 8%, 40-year-single payment present worth factor

Manufacturer *A* switch would be the overall lowest cost and would be the better deal provided the capability and reliability of the two switches are equivalent.

18.28 DISTRIBUTION SYSTEM LOSSES

About 8% of the total output of a large power system is lost or unaccounted for. Much of this loss is in the distribution system. Since investment must be made in facilities to supply these losses, they should be an important consideration in the engineering design of the system. A knowledge of their magnitude is essential and they should not be omitted from overall comparisons of alternative facilities without a study of each specific situation.

Line Losses. The line losses, which are the sum of the I^2R, or resistance losses, can be easily found when the currents at peak load are known. Simplifying assumptions often can be made in making these calculations. For instance, if the load can be considered as being uniformly distributed, the losses are the same as if the total load were concentrated at a point one-third of the way out on the feeder.

Transformer Losses. Transformers have a no-load loss as well as a load loss. The transformer no-load loss is independent of load, whereas the load loss will vary as the square of the current. These losses for distribution transformers are usually published as no-load and total loss when the transformer is operating at rated voltage and rated kVA. The load loss at full-load current is the difference between total and no-load losses.

Working Principles. The problem of converting kWh of lost energy to dollars and cents has resulted in considerable controversy among system operators because of the difficulty of determining the value of the energy. It is not the purpose of this handbook to take sides in the controversy but rather to show the principles involved so that engineers will be able to evaluate losses using appropriate system costs.

The cost of supplying losses can be broken down into two major parts:

1. Energy component, or production cost to generate kWh losses
2. Demand component, or annual costs associated with system investment required to supply the peak kW of loss

The two components of cost usually are combined into a single figure either in terms of cents per kilowatthour of total energy loss or as dollars per kilowatt of peak loss. Expressing losses in terms of dollars per kilowatt is usually called *capitalized* cost of losses, and it has some advantage in that it shows directly the amount of money that could be economically spent to save 1 kW of loss. However, the expression of cost of losses in cents per kilowatthour is usually a more convenient form to use in most engineering studies.

The cost of losses depends on the point in the system at which they occur. The farther out on the system, the greater value losses have. One kilowatt of loss saved on the secondary system is worth more than 1 kW loss at generation because of the cumulative effect of increments of losses as they pass through various elements of the system.

In calculating loss, present-day or future cost of system investment should be used. The primary interest is to find the incremental investment, in dollars, required to supply an incremental load in kilowatts.

Opinions differ widely as to the degree to which the demand component of losses shall be evaluated. This ranges all the way from the dollar cost per kilowatt for future system expansion to no value at all for this component. The great majority of utility engineers prefer to assign full value to the demand component of losses.

Responsibility Factor. Owing to diversity between classes of loads (i.e., residential, industrial, etc.) on a distribution system, peak loads on distribution, transmission, and generation usually do not

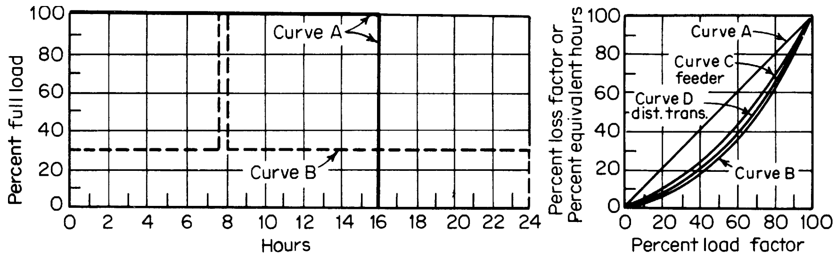


FIGURE 18-67 Relationship between load factor and loss factor or equivalent hours.

occur at the same time. Therefore, a loss which contributes 1 kW to the distribution system peak might contribute less than this to transmission and production plant peak because its maximum does not occur at the same time as the transmission or generation peak. This introduces *peak responsibility factors* used for evaluating cost of losses in various parts of the system.

Loss Factor. If the peak conductor losses of line or transformer have been calculated, it will still be necessary to know the loss factor or percent equivalent hours before it is possible to calculate the actual losses over a period of time. *Loss factor* is usually defined as the ratio of the average power loss, over a designated period of time, to the maximum loss occurring in that period. The term can refer to any part or all of the electric system. It is sometimes referred to as the *load factor of the losses*.

A corollary to loss factor is the term *equivalent hours*. This is defined as the number of hours per day, week, month, or year of peak load necessary to give the same total kilowatthours of loss as that produced by the actual variable load over the selected period of time. The period of time for distribution studies is usually 1 year, and it is obvious that *percent equivalent hours* has the same meaning as the term *percent loss factor*.

Relation Between Loss Factor and Load Factor. Definitions of *loss factor* and *load factor* are quite similar. (*Load factor* is defined as the ratio of average power demand over a stipulated period of time to the peak or maximum demand for that same interval.) Care should be taken that the latter is not used in place of loss factor when considering system losses. There is a relationship between the two factors which depends on the shape of the load curve. Because resistance losses vary as the square of the load, it can be shown that the value of loss factor can vary between the extreme limits of load factor and load factor squared. A number of typical load curves have been studied to determine this relationship for distribution feeders and distribution transformers. The relation is shown in Fig. 18-67. Note that loss factor is always less than load factor except where they are both unity, as would be the case for transformer core losses. The relationship between load factor and loss factor at the distribution transformer can be expressed by the empirical formula

$$\text{Loss factor} = 0.15 \text{ load factor} + 0.85 \text{ (load factor)}^2$$

It should be noted that when the shape of the load curve is known or can be reasonably estimated, the loss factor should be calculated directly and not determined by the empirical formula.

Cost of Losses. The two parts of the cost to supply losses are as follows:

$$\text{Energy component} = 8760 F_L E$$

$$\text{Demand component} = F_S P$$

where F_L = loss factor of load
E = cost of energy, dollars/kWh
F_S = responsibility factor
P = annual cost of system capacity, dollars/kW · year

Annual cost of losses can be combined into one value, in terms of either dollars per kilowatthour or dollars per kilowatt-year of peak loss, with the following formulas:

$$\text{Cost of losses, \$/kWh} = \frac{F_S P}{8760\, F_L} + E$$

$$\text{Capitalized cost of losses, \$/kW-year} = F_S P + 8760\, F_L E$$

18.29 STREET-LIGHTING SYSTEMS

Characteristics. The lighting of streets, parkways, and other roadways is about the only service in which the electric utility is often responsible for the utilization equipment. This involves the complete service of installation, maintenance, and operation of lighting systems during the hours of darkness (approximately 4000 h/yr) when they are required. Series (constant-current) circuits, historically a common supply for street lighting, have become obsolete.

Multiple Circuits. Street-lighting units today are normally supplied directly from the local distribution (120/240 V). The high-intensity-discharge lamps used have compatible balasts for all common voltages, 120, 208, 240, 277, 480 V, designed to strike an arc within the light source and provide stable operating conditions. Ballasts may be high-power-factor or normal-power-factor types. Photoelectric controls are most frequently used integrally with individual lights but also may be used to switch contactors controlling circuits used for lighting only. An example of this is highway lighting where extended systems from a power-supply point are normally designed with 480 V.

Lamps and Luminaires. Present street-lighting systems are designed using high-intensity-discharge sources. The three principal types are clear or phosphor-coated mercury, metal halide, and high-pressure sodium. These lamps are avaliable in several sizes ranging from less than 100 to 1000 W. Metal halide is not widely used because of its short life and poor lumen maintenance. Mercury lighting was the most popular type, since it was used extensively to replace older incandescent and fluorescent systems in the recent past. However, high-pressure sodium is the newest and most efficient lamp available. Its efficiency is over twice as high as mercury. The compact arc also allows for better control of the light distribution by the luminaire. The higher lamp efficiency and better control reduce the street-lighting power requirements to less than half that required for mercury. High-pressure sodium is taking over as the leading lighting system because of its economics.

Luminaires are sealed and also can be filtered. This minimizes the light loss due to dirt collecting in the luminaire. This is done to match the luminaire dirt depreciation so it matches the 4-year lamp life and minimizes the cleaning required during relamping.

There is a large variety of street-lighting equipment. This is required to fit different mounting heights, street widths, and lamp wattage. There are also differences in daytime appearance that are needed to fit the needs of the environment.

Underground Systems. While most utility-owned lighting systems were originally attached to existing wood-pole overhead distribution lines, these convenient supports are increasingly on rear-lot lines or nonexistent (underground distribution). This means that public lighting systems must be designed with underground supply run from the nearest transformer, joint trench, or handhole. Integrating the street-lighting circuits properly with the overall underground system from the outset is essential for proper economics and cost control.

18.30 RELIABILITY

Reliability has always been a major consideration for utilities. In recent years there has been even more interest because loads are becoming more sensitive to even small system disturbances and concern has been expressed that deregulation, with associated cuts in personnel and budgets, will negatively impact system reliability. Distribution reliability is in a state of change. Some of these changes are

- Standardization of indices.
- Mandatory indices in some states.
- Performance-based rates.
- Sags and momentaries equaling outages for some loads.
- Contract penalties for interruptions, sags, etc.
- Maintenance budgets being reduced.

There are many ways to measure reliability. Some of the more common indices used by utilities are CAIDI, SAIDI, and SAIFI, as illustrated in Fig. 18-68.

These three indices are defined for sustained interruptions of 5 minutes or longer:

SAIFI [*system average interruption frequency index (sustained interruptions)*]. The system average interruption frequency index is designed to give information about the average frequency of sustained interruptions per customer over a predefined area. In words, the definition is:

$$\text{SAIFI} = \frac{\text{total number of customer interruptions}}{\text{total number of customers served}}$$

To calculate the index, use the following equation:

$$\text{SAIFI} = \frac{\sum N_i}{N_T}$$

SAIDI (*system average interruption duration index*). This index is commonly referred to as Customer Minutes of Interruption or Customer Hours, and is designed to provide information about the average time the customers are interrupted. In words, the definition is:

$$\text{SAIDI} = \frac{\sum \text{customer interruption durations}}{\text{total number of customers served}}$$

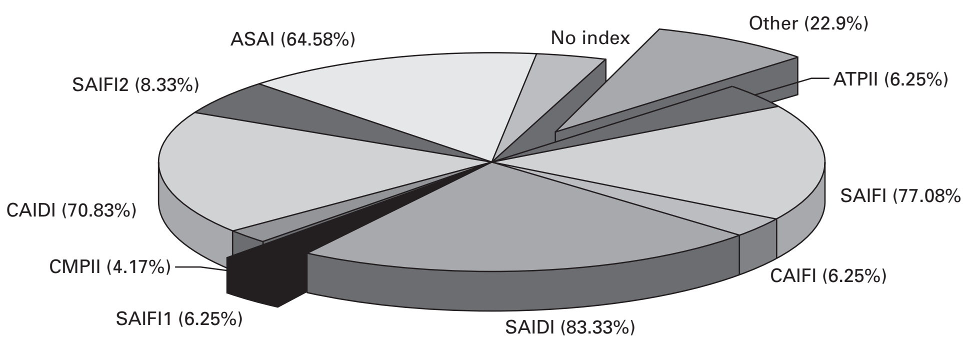


FIGURE 18-68 Percentage of companies using reliability indices (48 companies responded to survey).

To calculate the index, use the following equation:

$$\text{SAIDI} = \frac{\sum r_i N_i}{N_T}$$

CAIDI (*customer average interruption duration index*). CAIDI represents the average time required to restore service to the average customer per sustained interruption. In words, the definition is:

$$\text{CAIDI} = \frac{\sum \text{customer interruption durations}}{\text{total number of customers interruptions}}$$

To calculate the index, use the following equation:

$$\text{CAIDI} = \frac{\sum r_i N_i}{\sum N_i} = \frac{\text{SAIDI}}{\text{SAIFI}}$$

Values of these indices vary widely depending on many factors, including climate (snow, wind, lightning, etc.), system design (radial, looped, primary selective, secondary network, etc.), and load density (urban, suburban and rural). Typical values seen by utilities in the United States are

SAIDI	SAIFI	CAIDI
96 min/yr	1.2 int/yr	80 min/yr

Some utilities are already measuring indices to reflect system disturbances, other than interruptions, that cause sensitive loads to misoperate. One of these, the momentary average interruption-event frequency index (MAIFI_E), is an index to record momentary outages caused by successful reclosing operations of the feeder breaker or line recloser. This index is very similar to SAIFI, but it tracks the average frequency of momentary interruption events. In words, the definition is:

$$\text{MAIFI}_E = \frac{\text{total number of customer momentary interruption events}}{\text{total number of customers served}}$$

To calculate the index, use the following equation:

$$\text{MAIFI}_E = \frac{\sum \text{ID}_E N_i}{N_T}$$

Note. Here, N_i is the number of customers experiencing momentary interruptions events and ID_E equals interrupting device events during reporting period. This index does not include the events immediately preceding a lockout.

Another proposed index to reflect voltage sags caused by faults on other parts of the system is the system average rms (variation) frequency $\text{index}_{\text{Threshold}}$($\text{SARFI}_{\%V}$). This index records the number of specified short-duration rms variation per system customer. Voltage threshold allows assessment of compatibility for voltage-sensitive devices. To calculate the index, use the following equation:

$$\text{SARFI}_{\%V} = \frac{\sum N_i}{N_T}$$

where $\%V$ = rms voltage threshold 140, 120, 110, 90, 80, 70, 50, 10
N_i = number of customers experiencing rms $< \%V$ for variation i (rms $> \%V$ for $\%V > 100$)
N_T = total number of system customers

It is inevitable that more and more utilities will adopt some of these so-called power quality indices as their customers demand even better power for their sensitive loads.

In these days of reduced budgets, when utilities are being required to increase reliability, some of the techniques which cost very little or even nothing to achieve the goal of greater power quality are as follows:

- Purchase better-quality equipment.
- Shorten lead lengths on arresters.
- Use open-tie protection on underground systems.
- Use higher fuse ratings for transformers and laterals.
- Increase the number of homes per transformer.
- Pay attention to proper grounding.
- Use predictive reliability computer analysis to optimize designs.

18.31 EUROPEAN PRACTICES

In a time of deregulation and privatization, it has become common practice for a utility in one part of the world to own a utility in another country. While generation and transmission have relatively similar practices in all parts of the globe, distribution practices are considerably different depending on whether the system is based on American or European practices and standards. The following is a brief comparison of the two systems to familiarize engineers with the fact that in many ways distribution system operation and philosophy are so varied that direct comparison becomes extremely difficult.

Voltage Levels. In the United States, primary voltage levels can be just about anything. Figure 18-69 shows some of the more common voltage levels in the United States, with 13.8 kV probably being the most popular for the distribution primary. European voltage levels are much more standardized. Thus 30, 20, and 10 kV are used throughout the world where European standards are practiced.

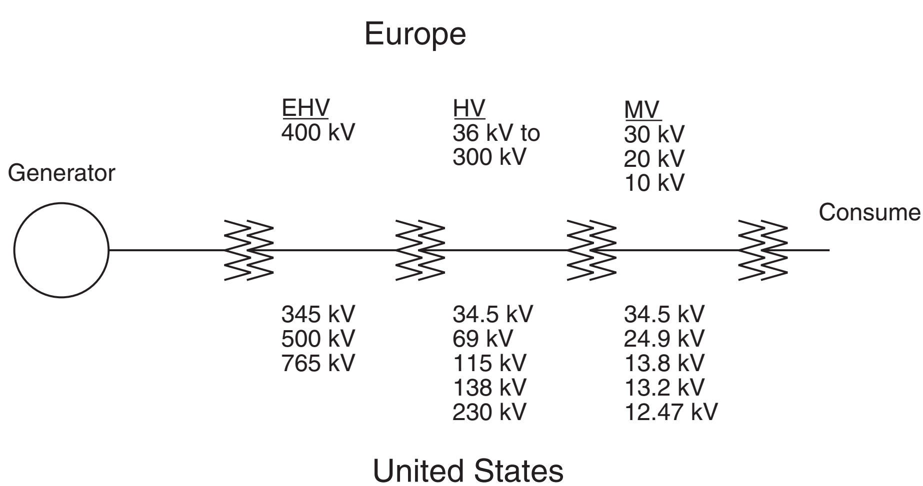


FIGURE 18-69 European and U.S. primary voltage levels.

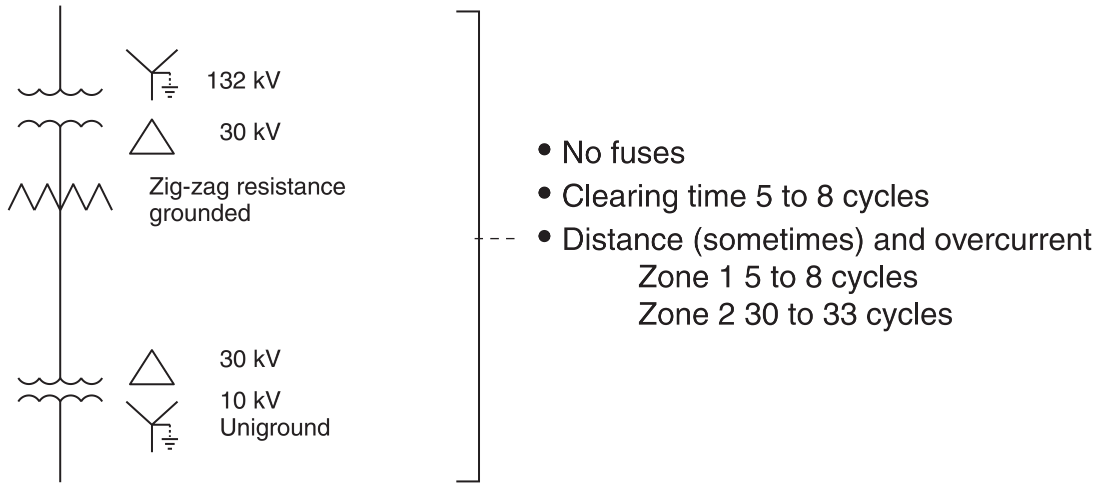


FIGURE 18-70 European distribution system grounding practices.

30-kV/10-kV Distribution. Figure 18-70 is a typical European system design, showing a ungrounded system at the 10-kV level where most of the distribution loads are supplied. This voltage level tends to be radial but can be networked. The 30-kV system, on the other hand, tends to be looped and is a 3-wire delta system sometimes using a grounding transformer to facilitate overcurrent protection.

Residential Distribution. Table 18-27 illustrates some of the major differences in philosophy between the two designs which makes comparison difficult. The major difference is that European practice tends to use 3-phase transformers with a much larger kVA rating to supply many more homes. Higher-density loads and higher secondary voltages are part of the reason for this difference.

Distributed Resources (DR) on the Distribution System. Distributed resources is a term which includes a variety of small generation technologies, including fuel cells, photovoltaics, microturbines, reciprocating engines, wind turbines, etc., with and without battery storage, which could be installed on the distribution system. In addition, many of these resources are becoming more

TABLE 18-27 1-Phase vs. 3-Phase Residential Distribution

United States	Europe
120/240 V	400 Wye/230, 4-wire (Europe)
Single-phase transformers heavily overloaded—25 kVA typical	Less load per home than U.S.
Four homes/transformer fairly typical	3-Phase xfrms >> 1-phase
Fuses are typically expulsion	Residential units in 300–500-kVA range, 5 to 10 radial 3-phase 4-wire secondary feeds, per transformer
	No overload Fuses are current-limiting 100 to 200 dwellings per transformer

modular in design, so that manufacturing economies of scale are driving their costs down. Generally, distributed resources are small in size, ranging from less than 1 kW to a few hundred kWs. The practical size limit for generators on the distribution system is in the area of around 35 MW.

The justification for introducing these newer technologies include the renewable resource aspect and lower environmental impact along with niche opportunities in particular regions when one or more of these technologies might excel. It is claimed that more DR means less T&D investment. Without storage, many DR technologies provide an additional energy source but no demand or investment savings. Then too improved reliability is often touted as a benefit, but this needs to be property evaluated in the context of what technology is being discussed. Without careful engineering, a particular MW level of DR penetration can adversely affect the distribution system.

Distribution designs are based on a number of principles that can be upset by DR including:

- Most distribution circuits are radial in nature with power flow from the substation to the loads.
- Voltage control with line drop compensation through LTC transformers or voltage regulators presumes a dropoff of load and voltage as one proceeds from the substation and, further, that power flow is from the source side to the load side.
- Most overcurrent protective device selection and coordination is based on higher fault current magnitudes at the substation and declining fault current magnitudes toward the end of the circuit.
- Reclosing intervals of circuit breakers and reclosers and reclosing practice in general is based on reasonable fault clearing times and the expectation that the reclosing device will close into a de-energized line.
- Utility restoration practices are based on having a limited number of known sources controlled by switches and other protective devices with known states (energized or de-energized).
- Harmonics are limited to larger known sources and distributed smaller sources.
- Most 3-phase transformer connections on distribution are grounded-wye grounded-wye.

With the arrival of distribution resources, many of these *normal conditions* now change, including:

- Bidirectional power flow and fault current flow along portions of the circuit.
- Load magnitude dropping off along the feeder and then reaching a step change at the DR location(s) and dropping off beyond that point.
- Fault current profile showing the increase in total fault current due to added generation as well as the bidirectional fault current flows from the DRs depending on fault location.
- Automatic reclosing for utility circuit breakers and reclosers will have to be supervised with some form of voltage check and possibly time delay. An alternative is to use transfer-tripping schemes to assure that the DR is off-line before reclosing takes place.
- Many DR advocates seek a delta-wye connection. Under backfeed conditions, a single-line-to-ground fault on the primary will overstress the surge arresters on the unfaulted phases.
- Photovoltaics and many small wind generators are dc machines that rely on invertors to produce an ac waveform and also produce high harmonic content. Induction-type machines draw reactive power from the power system. The combination introduces both voltage drop and power quality concerns.

IEEE Std 1547-2003, Standard for Interconnecting Distributed Resources with Electric Power Systems, was developed to provide a uniform standard for the interconnection of distributed resources with electric power systems. It provides requirements relevant to the performance, operation, testing, safety considerations, and maintenance of the interconnection. Follow-on standards projects to Std 1547 are intended to address the different DR technologies and engineering concerns about the interconnection issues.

A logical concern is—at what level of distributed resources does one have to be concerned with some of these potential issues becoming genuine problems? Unfortunately, we don't yet have an answer to that question.

BIBLIOGRAPHY

Books for General Reference

Blume, L. F., Boyajian, A., Camilli, G., Lennox, T. C., Minneci, S., and Montsinger, V. M.: *Transformer Engineering*, Wiley, 1951.

Burke, James J.: *Power Distribution Engineering; Fundamentals and Applications*, Marcel Dekker, 1994.

Edison Electric Institute: *Underground Systems Reference Book*, Edison Electric Institute, 1957.

Grainger, J. J. and Stevenson, W. D.: *Power System Analysis*, McGraw-Hill, 1994.

Greenwood, A.: *Electrical Transients in Power Systems*, Wiley-Interscience, 1971.

Lewis, W. W.: *Protection of Transmission Systems against Lightning*, Wiley, 1950.

Pansini, A. I.: *Electrical Distribution Engineering,* McGraw-Hill, 1983.

Peterson, H. A.: *Transients in Power Systems,* Dover Publications, 1951.

Short, T. A.: *Electrical Power Distribution Handbook*, CRC Press, 2004.

Thue, W. A.: *Electrical Power Cable Engineering,* Marcel Dekker, 1999.

Manufacturers' Publications

Distribution Data Book, GET-1008M, General Electric Company, 1980.

Distribution System Feeder Overcurrent Protection, GET-6450, General Electric Company, 1977.

Distribution Transformer Guide, ABB, October 1991.

Electric Utility Engineering Reference Book—Distribution Systems, Westinghouse Electric Corporation, vol. 3, 1965.

Electrical Distribution—System Protection, Cooper Power Systems, 3rd ed., 1990.

Electrical Transmission and Distribution Reference Book, Westinghouse Electric Corporation, 1964.

Getting Down to Earth, Biddle Instruments, 1982.

Overhead Conductor Manual, Southwire Company, 1994.

Power Cable Manual, Southwire Company, 1997.

IEEE Transactions (Formerly AIEE)

System

Bankus, H. M. and Gerngross, J. E.: Unbalanced Loading and Voltage Unbalance on Three-Phase Distribution Transformer Banks, 1954, vol. 73, pt. III, p. 367.

Bankus, H. M. and Gerngross, J. E.: Combined Single-Phase and Three-Phase Loading of Open-Delta Transformer Banks, *Power Apparatus and Systems*, February 1958, pp. 1337–1343.

Easley, J. H. and Shula, W. E.: Cost and Reliability Evaluation of Four Underground Primary Distribution Feeder Plans, transactions paper, Conference Record—1974 Underground Transmission and Distribution Conference, 74-CH0832-6-PWR, pp. 436–443.

Mitchell, C. F., Sweeney, J. O., and Cantwell, J. L.: An Economic Analysis of Distribution Transformer Application, *Power Apparatus and Systems*, December 1959, pp. 1196–1202.

Nickel, D. L.: Distribution Transformer Loss Evaluation. I—Proposed Techniques, *Power Apparatus and Systems*, vol. PAS-100, no. 2, February 1981, pp. 788–97.

Ward, D. J., Griffith, D. C., and Burke, J. J.: Power Quality—Two Different Perspectives, *Trans. on Power Delivery*, vol. 5, no. 3, July 1990, pp. 1501–13.

System Planning

Anderson, A. S. and Thiemann, V. A.: Distribution Secondary Conductor Economics, *Power Apparatus and Systems*, February 1960, pp. 1839–1843.

Blake, D. K.: Some Observations on the Economic Benefits in Going from One System Voltage Level to a Higher System Voltage Level, *Power Apparatus and Systems*, vol. 71, pt. III, pp. 585–592.

Campbell, H. E., Ender, R. C., Gangel, M. W., and Talley, V. C.: Economic Analysis of Distribution Systems, *Power Apparatus and Systems*, August 1960, pp. 423–443.

Jones, A. I., Smith, B. E., and Ward, D. J.: Considerations for Higher Voltage Distribution, *Trans. on Power Delivery*, April 1992, pp. 782–788.

Rudasill, C. L. and Ward, D. J.: Distribution Underground Cable Evaluation, *Trans. on Power Delivery*, July 1997, vol. 12, no. 3, pp. 1398–1403.

Sarkas, R. H. and Thacker, H. B.: Distribution System Load Characteristics and Their Use in Planning and Design, *Power Apparatus and Systems*, August 1957, pp. 564–573.

Schultz, N. R.: Distribution Primary Feeder I^2R Losses, *Power Apparatus and Systems*, March/April 1978, vol. PAS-97, no. 2, pp. 603–9.

Smith, J. A.: Determination of Economical Distribution Substation Size, *Power Apparatus and Systems*, October 1961, pp. 663–670.

Smith, J. A.: Economics of Primary Distribution Voltages of 4.16 through 34.5 kV, *Power Apparatus and Systems*, October 1961, pp. 670–683.

Van Wormer, F. C.: Some Aspects of Distribution Load Area Geometry, *Power Apparatus and Systems*, December 1954, pp. 1343–1349.

Webler, R. M., Gangel, M. W., Carter, G. K., Zeman, A. L., and Ender, R. C.: Secondary Distribution System Planning for Load Growth, *Power Apparatus and Systems*, December 1963, pp. 908–927.

Overvoltage and Overvoltage Protection

Burke, J. J., Sakshaug, E. C., and Smith, S. L.: The Application of Gapless Arresters on Underground Distribution Systems, *Trans. on Power Apparatus and Systems*, March 1981, vol. 100, pp. 1234–1243.

Clayton, J. M. and Hileman, A. R.: A Method of Estimating Lightning Performance of Distribution Lines, *Power Apparatus and Systems*, 1954, vol. 73, pt. III, p. 953.

Headrickson, P. E., Johnson, I. B., and Schultz, N. R.: Abnormal Voltage Conditions Produced by Open Conductors on Three-Phase Circuits Using Shunt Capacitors, *Power Apparatus and Systems*, 1953, vol. 72, pt. III, p. 1183.

Hopkinson, R. H.: Better Surge Protection Extends URD Cable Life, *Trans. on Power Apparatus and Systems*, October 1984, vol. 103, pp. 2827–2836.

Hopkinson, R. H.: Ferroresonance during Single-Phase Switching of Three-Phase Distribution Transformer Banks, *Power Apparatus and Systems*, Apri1 1965, vol. PAS-4, pp. 289–293. *Discussion*, June 1965, pp. 514– 517.

Kershaw, S. S., Gaibrois, G. L., and Stump, K. B.: Applying Metal Oxide Surge Arresters on Distribution Systems, *Trans. on Power Delivery*, January 1989, vol. 4, no.1, pp. 301–307.

Mancao, R. T., Short, T. A., and Burke, J. J.: Application of MOVs in the Distribution Environment, *Trans. on Power Delivery*, January 1994, vol. 9, no. 1, pp. 293–305.

Sakshaug, E. C., Kresge, J. S., and Miske, S. A., Jr.: A New Concept in Station Arrester Design, *Trans. on Power Apparatus and Systems*, March/April 1977, vol. 96, p. 647.

Short, T. A.: Distribution Lightning Performance Calculations, *IEEE Computer Applications in Power*, November 1991.

Task Force Report-Investigation and Evaluation of Lightning Protective Methods for Distribution Circuits—Part I, Model Study and Analysis: Part II, Applications and Evaluation, *Power Apparatus and Systems*, August 1969, vol. PAS-88, no.8, pp. 1232–1247.

Working Group of Surge Protective Devices Committee: Voltage Rating Investigation for Application of Lightning Arresters on Distribution Systems, *Trans. on Power Apparatus and Systems*, May/June 1972, vol. 91, no.3, pp. 1067–1074.

Overcurrent and Overcurrent Protection

Arndt, R. H., Koch, R. E., and Schultz, N. R.: Concept Alternatives and Application Considerations in the Use of Current-Limiting Fuses for Transformer Protection, Transactions Paper, Conference Record—1974 Underground Transmission and Distribution Conference, 74-CH0832-6-PWR, pp. 259–267.

Auer, G. G., Ender, R. C., and Wylie, R. A.: Digital Calculation of Sequence Impedances and Fault Currents for Radial Primary Distribution Circuits, *Power Apparatus and Systems*, February 1961, pp. 1264–1277.

Burke, J. J. and Lawrence, D. J.: Characteristics of Fault Currents on Distribution Systems, January 1984, PAS vol. 103, no. 1, pp. 1–6.

Harner, R. H.: Secondary-Fault Recovery Voltage Investigation, *Power Apparatus and Systems*, February 1968, vol. PAS-87, no.2, pp. 463–487.

IEEE Tutorial Course on Application and Coordination of Reclosers, Sectionalizers and Fuses, Publication 80 EHO157-8-PWR, 1980.

Voltage Regulation and Kilovar Supply

Barger, J. V. and Smith, D. R.: Impedance and Circulating Current Calculations for UD Multi-Wire Neutral Circuits, *Power Apparatus and Systems*, May–June 1972, vol. PAS-91, no.3, pp. 992–1006.

Grainger, J. J. and Lee, S. H.: Optimum Size and Location of Shunt Capacitors for Reduction of Losses on Distribution Feeders, *Power Apparatus and Systems*, March 1981, vol. PAS-100, no. 3, pp. 1105–18.

Johnson, I. B., Schultz, A. J., Schultz, N. R., and Shores, R. B.: Some Fundamentals on Capacitance Switching, *Power Apparatus and Systems*, August 1955, pp. 727–736.

Neagle, N. M. and Samson, D. R.: Loss Reduction from Capacitors Installed on Primary Feeders, *Power Apparatus and Systems*, October 1956, pp. 950–959.

Grounding

Application Guide and Methods of Substation Grounding, AIEE Group on Substation Grounding Practices, 1954, vol. 73, pt. m, p. 271.

Mancao, R. T., Myers, A., and Burke, J. J.: The Effect of Distribution System Grounding on MOV Selection, *Trans. on Power Delivery,* January 1993, vol. 8, **1**.

Standards and Standards Publications

American National Standards Institute (ANSI) C2-2002, National Electrical Safety Code.

ANSI C84.1-1995 (R2001), Voltage Ratings for Electric Power Systems and Equipment (60-Hz).

ANSI/ICEA S-94-649 Concentric Neutral Cables Rated 5-46 kV, 2000.

ANSI/ICEA S-97-682 Utility Shielded Power Cable Rated 5-46 kV, 2000.

IEEE Std 835-1994, Power Cable Ampacity Tables.

NEC 2005, NFPA 70, National Electric Code.

Periodicals

System

Burke, J. J.: Utility Characteristics Affecting Sensitive Industrial Loads, *Power Quality Assurance Magazine*, November–December 1996.

Gangel, M. W. and Propst, R. F.: Investigating Distribution Transformer Load Characteristics, *Distribution Magazine*, July 1961, p. 6.

System Planning

Brown, P. G., Propst, H. R., and Tice, J B.: Unity Power Factor Is Essential to Emergency Kilowatt Transportation, *Electric Forum Magazine*, Fall 1975, p. 10.

Campbell, H. E.: Serving Critical Loads, *Distribution Magazine*, 1966, 4th quarter, p. 9.

Hayes, R. H. and Hill, 0. L.: Progress in Remote Line Switch Control, *Transmission and Distribution*, June 1975, p. 52.

Van Wormer, F. C.: Design and Operation of Spot Networks, *Distribution Magazine*, 1966, 2d/3d quarter p. 5; 1966, 4th quarter, p. 19.

Overvoltage and Overvoltage Protection

Auer, G. G.: Basic Considerations in Lighting Protection of URD Systems, *Distribution Magazine*, April 1968, p. 16.

Barker, P. P. and Burke, J. J.: Protecting Underground Distribution Systems, *Electric Light and Power*, Apri1 199l.

Overcurrent and Overcurrent Protection

Howard, S.B. and Stroebed, R. W.: Can Single-Phase Cutouts Be Applied to Three-Phase Circuits, *Distribution Magazine*, 1964, 2d quarter, p. 4.

Lasseter, J. A.: Burndown Tests on Bare Conductors, *Electric Light and Power*, December 15, 1956, p. 94.

Voltage Regulation and Kilovar Supply

Gangel, M. W.: Compensator Settings Made Easier, *Distribution Magazine*, pt. 1, April 1960, p. 22; pt. 2, July 1960, p. 18.

Schultz, N. R.: Calculating Voltage Drop and Power Loss, *Distribution Magazine*, January 1969, p. 11.

Underground Systems

Van Wormer, F. C.: Underground Distribution Systems for Residential Areas, *Distribution Magazine*, 1, January 1959, p. 3; pt. 2, Apri1 1959, p. 12; pt. 3, April 1960, p. 16; pt. 4, Apri1 1962, p. 3; pt. 5, Apri1 1963, p. 22.

Miscellaneous Publications

System

Beaty, H. Wayne: 10th Annual T&D Construction Survey, Electrical World, September 1, 1975, pp. 35–42.

Dudas, J. and Fletcher, C.: Underground Cable Specification Advances and Installation Practices of the Largest Investor Owned Utilities, Fall Insulated Conductors Committee Meeting, 2004.

Gangel, M. W. and Propst, R. F.: Transformer Characteristics Correlated to Loading: Power Distribution Conference, University of Texas, October 1963.

RUS Specifications and Drawings for 12.5/7.2 kV Line Construction 5/83, Bulletin 50-3, 2005.

RUS Specification and Drawings for 34.5/19.9 kV Distribution Line Construction (11–86), Bulletin 50-4.

RUS Specifications and Drawings for Underground Electric Distribution (3–90), Bulletin 50-6, 2000.

RUS Specifications and Drawings for Underground Cable Installation, Doc. 345-152, Form 515d, 1989.

System Planning

Campbell, H. E.: Today and Tomorrow, Underground Distribution to High Rise Buildings, IEEE Conference Record-Special Technical Conference on Underground Distribution, 31C35, September 1966, pp. 223–239.

Crawford, J. W. and Hamner, F. G.: Demand and Diversity Characteristics of Residential Loads, Southeastern Electric Exchange, Engineering and Operating Conference, Apri1 1963.

Load Growth Forces Higher Voltages, Electrical World, June 1, 1974, pp. 154–163.

Overcurrent Protection

Beaty, H. Wayne: Special Report-Switching and Overcurrent Protection for Distribution Systems, Electrical World, April 1, 1974, pp. 41–56.

Campbell, H. E.: Implication of Increased Short-Circuit Duty on Residential Distribution Systems, American Power Conference, vol. 35, 1973, pp. 1098–1104. Underground Systems.

IEEE Conference Record-1974 Underground Transmission and Distribution, 74CHO832-6-PWR and 74CHO832-6-PWR (SUP.), April 1–5, 1974.

IEEE Conference Record-1991 Transmission and Distribution Conference, 911CH3070-0, September 1991.

Lewis, S. M.: URD Survey Report, Transmission and Distribution, July 1973, pp. 88–95.

Specifications and Drawings for Underground Electric Distribution, RUS Bulletin 50-6, Rural Utilities Service, U.S. Department of Agriculture, March 1990.

Underground Corrosion Control Guide, NRECA Research Project, August 1982.

SECTION 19
WIRING DESIGN FOR COMMERCIAL AND INDUSTRIAL BUILDINGS

John Dagenhart, P. E.
Professional Engineer, Clapp Research Associates. P. C.

CONTENTS

19.1 BASIC INSTALLATION RULES AND INSPECTION

Codes (Definitions). The *National Electrical Code* (NEC)*
The *National Electrical Safety Code* (NESC) establishes the basic standards of electric supply system design and installation for utility-owned conductors and equipment in the United States. It is also revised periodically by a committee drawn from utility groups, industries, state and federal regulators, insurance groups, organized labor, and other interested parties. Its secretariat is the Institute of Electrical and Electronics Engineers; the NESC is American National Standard ANSI C2. The NEC oversees supply and communication wiring that are in and on consumer-owned buildings but not an integral part of a generating plant, substation, or control center. The NEC does not cover communication utility wiring, nor does it cover electric utility generation, transmission, or distribution system wiring. The NESC covers the latter systems. The NESC also covers similar systems under the control of qualified persons, such as those associated with large industrial complexes. In recent years, the provisions of the NESC relating to underground wiring have become increasingly applicable in commercial complexes as extremely large commercial complexes have become more frequent. Some of the latter systems are not unlike those utility systems found in small towns or compact subdivisions.

Lists of Inspected Electrical Equipment and Appliances are issued yearly by the *Underwriters' Laboratories, Inc. Electrical Testing Laboratories, Inc., Factory Mutual Research Corp.*, and *MET Electrical Testing Company, Inc*. are other testing laboratories that function as third-party certifiers of the basic safety of manufactured products used in electrical work. One function of the laboratories

**National Electrical Code* and the acronym *NEC* are registered trademarks of the National Fire Protection Association, Inc., Quincy, Mass. 02269.

Establishes the standards of wiring design and installation practice for consumer-owned wiring and equipment in the United States. Its rules are written to protect the public from fire and life hazards. It is revised periodically by a committee drawn from industry associations, insurance groups, organized labor, and representatives of municipalities. It is sponsored by the National Fire Protection Association, and approved by the American National Standards Institute as ANSI C1. It forms the basis of the vast majority of municipal electrical wiring ordinances, which adopt successive editions of the Code as issued.

is to examine and pass on electrical materials, fittings, and appliances in order to determine if they comply with the standard-test specifications set up by these laboratories.

Legal Status of the Code. The rules in the NEC are enforced by being incorporated in ordinances passed by various cities and towns, covering the installation of electric wiring. The Occupational Safety and Health Administration (OSHA) requires that all new electrical installations conform to all the rules of the NEC. The NESC is adopted by state utilities commissions and is referred to by the NEC for some high-voltage applications.

When installing any electrical equipment, first ascertain whether local installation rules in the form of ordinances are enforced in the community. If so, follow such rules; if none exists, follow the requirements of NEC.

Editions. Where reference is made in this section to installation rules, the 1996 edition of NEC or 1997 NESC is used as a basis.

Code Not a Design Manual. Design of an installation in accordance with the Code minimizes fire and accident hazards but does not guarantee satisfactory or efficient operation of the system. Other design standards are necessary to accomplish the latter purposes.

License. In many areas the installation of electric wiring is controlled by city, county, or state license, often combined with installation rules.

Rules of Electric Service Companies. Electric lighting and power companies generally issue certain rules of their own, based to a large extent on peculiar requirements which are necessary in order to give the best possible service to the greatest number of customers and on NESC requirements.

These rules are concerned mostly with matters of distribution engineering. They relate to locations and details of service entrance, provision for meters, the kind of electricity furnished by the company, its frequency and voltage, the types and sizes of motors, rules in connection with starting characteristics of such motors, and similar matters.

The electric-service company usually supplies copies of its rules at no charge.

Inspection. Every electrical installation should be inspected wherever an experienced inspector is available to ensure that it complies with local and NEC rules. Such inspection is usually mandatory in cities having electrical ordinances. In some areas the fire underwriters maintain inspectors who check electrical wiring, while in others the municipality makes a check through its electrical inspectors. Where inspection is not mandatory, it is always advisable to request the most convenient fire underwriters' bureau to make the necessary inspection.

Federal and state buildings usually require inspection by authorized federal or state inspectors. In these instances inspection includes not only safety considerations but the requirements of the particular job specifications. Other inspection may be required but it is often waived. OSHA compliance officers do make inspections of existing electrical systems at any time.

19.2 METHODS OF WIRING

Wiring Methods Classified. The discussion of wiring methods in this section relates to *interior circuits for light, heat, and power* and does not cover signaling or communication systems.

Numerous methods of wiring are authorized by NEC, most of them used to a greater or lesser extent in commercial and industrial buildings. Those of interest can be grouped as follows:

1. Raceways for general use

- ***a.*** Rigid-metal conduit
- ***b.*** Intermediate-metal conduit (IMC)
- ***c.*** Electric-metallic tubing (EMT)

   - ***d.*** Nonmetallic conduit
   - ***e.*** Surface raceways
   - ***f.*** Flexible metallic and nonmetallic conduit
   - ***g.*** Gutters

2. Cable-assembly systems for general use
   - ***a.*** Nonmetallic sheathed cable
   - ***b.*** Underground feeder and branch-circuit cable
   - ***c.*** Metal-clad cable (armored cable)
   - ***d.*** Mineral-insulated metal-sheathed cable (MIMS)
   - ***e.*** Messenger-supported wiring
   - ***f.*** Nonmetallic-sheathed cable (NM, NMC, NMS)
   - ***g.*** Power and control cable (TC)
   - ***h.*** Armored cable
3. Conductor systems for general use
   - ***a.*** Open wiring on insulators
   - ***b.*** Concealed knob and tube wiring (only as permitted in NEC Sec. 394)
4. Cable-assembly systems for limited use
   - ***a.*** Service-entrance cable
   - ***b.*** Nonmetallic extensions
   - ***c.*** Integrated gas spacer cable (IGS)
   - ***d.*** Medium-voltage cable (MV)
   - ***e.*** Flat conductor cable (FCC)
5. Raceway systems for limited use
   - ***a.*** Flexible-metal conduit and flexible-metal tubing
   - ***b.*** Liquidtight-flexible-metal conduit and liquidtight flexible nonmetallic conduit
   - ***c.*** Underfloor raceway
   - ***d.*** Cellular-metal-floor or cellular-concrete-floor raceway
   - ***e.*** Wireways
   - ***f.*** Cable trays
6. Special systems
   - ***a.*** Busways
   - ***b.*** Cable bus
   - ***c.*** Multioutlet assemblies
   - ***d.*** Electrical floor assemblies
   - ***e.*** Flat cable assemblies

Installation Methods. Requirements to be met in installing each of the foregoing systems are found in the current edition of the NEC. The requirements are specific and detailed and change somewhat as the art progresses; hence reference should be made to the Code for the exact circumstances under which each system is permitted or prohibited, together with the precise rules to be followed in installation.

The discussion in the following paragraphs compares the systems generally and indicates the major limitations on use of each.

19.3 TYPES OF CONDUCTOR

General Provisions Applying to All Wiring Systems. The types of wiring discussed may be used for voltages up to 2000 V unless otherwise indicated. Each type of insulated conductor is approved for certain uses and has a maximum operating temperature. If this is exceeded, the insulation is subject to deterioration. In recent years, conductors with asbestos insulation, formerly used for high-temperature operations, have been removed from the tables of conductor applications and insulations (see Table 19-1). Each conductor size has a maximum current-carrying capacity, depending on type of insulation and conditions of use. These ratings should not be exceeded (see Tables 19-2A through 19-2E and Fig. 19-1 for ratings and underground conduit systems). Conductors may be used in multiple usually in large sizes only (sizes 1/0 and larger, see NEC Sec. 310-4).

Conductors of more than 600 V should not occupy the same enclosure as conductors carrying less than 600 V, but conductors of different light and power systems of less than 600 V may be grouped together in one enclosure if all are insulated for the maximum voltage encountered. In general, communication circuits should not occupy the same enclosure with light and power wiring.

Boxes or fittings must be installed at all outlets, at switch or junction points of raceway or cable systems, and at each outlet and switch point of concealed knob and tube work.

Provisions Applying to All Raceway Systems. The number of conductors permitted in each size and type of raceway is definitely limited to provide ready installation and withdrawal. For conduit and EMT, see Table 19-3. Raceways, except surface-metal molding, must be installed as complete empty systems, the conductors being drawn in later. Conductors must be continuous from outlet to outlet without splice, except in auxiliary gutters and wireways.

Conductors of No. 8 American wire gauge (AWG) and larger must be stranded. Raceways must be continuous from outlet to outlet and from fitting to fitting and shall be securely fastened in place. Conductors and cables exposed to the sun must be sunlight resistant (see NEC Article 310.8(D)).

All conductors of a circuit operating on ac, if in metallic raceway, should be run in one enclosure to avoid inductive overheating. If, owing to capacity, not all conductors can be installed in one enclosure, each raceway used should contain a complete circuit (one conductor from each phase).

Rigid-Metal Conduit, Intermediate-Metal Conduit, and Electrical Metallic Tubing. These systems are systems generally employed where wires are to be installed in raceways. Both conduit and tubing may be buried in concrete fills or may be installed exposed. Wiring installed in conduit is approved for practically all classes of buildings and for voltages both above and below 600 V. Certain restrictions are placed on the use of tubing.

Metal conduit consists of standard-weight steel pipe (preferably either galvanized or cadmium-plated, although it may be black-enameled for use indoors and where not subject to severe corrosive influences) or of aluminum. Electrical metallic tubing has the same internal diameter as conduit but a thinner wall of higher-quality steel.

Note on Tables 19-2A through 19-2E: Use of conductors with higher operating temperatures. Where the room temperature is within 10°C of the maximum allowable operating temperature of the insulation, it is desirable to use an insulation with a higher maximum allowable operating temperature.

Fittings and connectors used with conduit may be threaded or threadless. Electrical metallic tubing fittings are usually threadless.

Nonmetallic rigid conduits, in approximately the same dimensions as rigid-metal conduits, are also a general-use raceway. Some restrictions are imposed, affecting particularly installations exposed to possible mechanical injury. Grounding continuity is provided by an additional grounding conductor pulled into the raceway with the circuit conductors or as part of a cable assembly.

Nonmetallic polyvinyl chloride (PVC) rigid conduits are commonly assembled with matching fittings by adhesives. Field bends are made by softening the plastic in a hot airstream of several hundred degrees from an electric heater-blower.

Nonmetallic PVC raceways of relatively flexible construction and with conductor already drawn in are used for direct burial in airport, highway, parkway, and similar installations.

Polyvinyl chloride and fiber conduits are extensively used in underground distribution. They may be installed directly in earth or encased in concrete envelopes.

TABLE 19-1 Conductor Application and Insulations

Trade Name	Type Letter	Maximum Operating Temperature	Application Provisions	Insulation	Thickness of insulation AWG or kcmil	mm	mils	Outer Covering[a]
Fluorinated ethylene propylene	FEP or FEPB	90°C 194°F	Dry and damp locations	Fluorinated ethylene propylene	14–10 8–2	0.51 0.76	20 30	None
		200°C 392°F	Dry locations—special applications[b]	Fluorinated ethylene propylene	14–8	0.36	14	Glass braid
					6–2	0.36	14	Glass or other suitable braid material
Mineral insulation (metal sheathed)	MI	90°C 194°F	Dry and wet locations	Magnesium oxide	18–16[c] 16–10 9–4 3–500	0.58 0.91 1.27 1.40	23 36 50 55	Copper or alloy steel
		250°C 482°F	For special applications[b]					
Moisture-, heat-, and oil-resistant thermoplastic	MTW	60°C 140°F	Machine tool wiring in wet locations	Flame-retardant, moisture-, heat-, and oil-resistant thermoplastic		(A) (B)	(A) (B)	(A) None (B) Nylon jacket or equivalent
		90°C 194°F	Machine tool wiring in dry locations FPN: See NFPA 79.		22–12 10 8 6 4–2 1–4/0 213–500 501–1000	0.76 0.38 0.76 0.51 1.14 0.76 1.52 0.76 1.52 1.02 2.03 1.27 2.41 1.52 2.79 1.78	30 15 30 20 45 30 60 30 60 40 80 50 95 60 110 70	
Paper		85°C 185°F	For underground service conductors, or by special permission	Paper				Lead sheath
Perfluoro alkoxy	PFA	90°C 194°F	Dry and damp locations	Perfluoro alkoxy	14–10 8–2 1–4/0	0.51 0.76 1.14	20 30 45	None
		200°C 392°F	Dry locations—special applications[b]					
Perfluoro alkoxy	PFAH	250°C 482°F	Dry locations only. Only for leads within apparatus or within raceways connected to apparatus (nickel or nickel-coated copper only)	Perfluoro alkoxy	14–10 8–2 1–4/0	0.51 0.76 1.14	20 30 45	None
Thermoset	RHH	90°C 194°F	Dry and damp locations		14–10 8–2 1–4/0 213–500 501–1000 1001–2000 For 601–2000 see Table 310.62.	1.14 152 2.03 2.41 2.79 3.18	45 60 80 95 110 125	Moisture-resistant, flame-retardant, nonmetallic covering[a]
Moisture-resistant thermoset	RHW[d]	75°C 167°F	Dry and wet locations	Flame-retardant, moisture-resistant thermoset	14–10 8–2 1–4/0 213–500 501–1000 1001–2000 For 601–2000 see Table 310.62.	1.14 1.52 2.03 2.41 2.79 3.18	45 60 80 95 110 125	Moisture-resistant, flame-retardant, nonmetallic covering[e]

(*Continued*)

TABLE 19-1 Conductor Application and Insulations (*Continued*)

Trade Name	Type Letter	Maximum Operating Temperature	Application Provisions	Insulation	Thickness of insulation			
					AWG or kcmil	mm	mils	Outer Covering[a]
Moisture-resistant thermoset	RHW-2	90°C 194°F	Dry and wet locations	Flame-retardant, moisture-resistant thermoset	14–10 8–2 1–4/0 213–500 501–1000 1001–2000 For 601–2000, see Table 310.62.	1.14 1.52 2.03 2.41 2.79 3.18	45 60 80 95 110 125	Moisture-resistant, flame-retardant, nonmetallic covering[e]
Silicone	SA	90°C 194°F 200°C 392°F	Dry and damp locations For special application[b]	Silicone rubber	14–10 8–2 1–4/0 213–500 501–1000 1001–2000	1.14 1.52 2.03 2.41 2.79 3.18	45 60 80 95 110 125	Glass or other suitable braid material
Thermoset	SJS	90°C 194°F	Switchboard wiring only	Flame-retardant thermoset	14–10 8–2 1–4/0	0.76 1.14 2.41	30 45 95	None
Thermoplastic and fibrous outer braid	TBS	90°C 194°F	Switchboard wiring only	Thermoplastic	14–10 8 6–2 1–4/0	0.76 1.14 1.52 2.03	30 45 60 80	Flame-retardant, nonmetallic covering
Extended polytetra-fluro-ethylene	TFE	250°C 482°F	Dry locations only. Only for leads within apparatus or within raceways connected to apparatus, or as open wiring (nickel or nickel-coated copper only)	Extruded polytetra-fluoro-ethylene	14–10 8–2 1–4/0	0.51 0.76 1.14	20 30 45	None
Heat-resistant thermoplastic	THHN	90°C 194°F	Dry and damp locations	Flame-retardant, heat-resistant thermoplastic	14–12 10 8–6 4–2 1–4/0 250–500 501–1000	0.38 0.51 0.76 1.02 1.27 1.52 1.78	15 20 30 40 50 60 70	Nylon jacket or equivalent
Moisture- and heat-resistant thermoplastic	THHW	75°C 167°F 90°C 194°F	Wet location Dry location	Flame-retardant, moisture- and heat-resistant thermoplastic	14–10 8 6–2 1–4/0 213–500 501–1000	0.76 1.14 1.52 2.03 2.41 2.79	30 45 60 80 95 110	None
Moisture- and heat-resistant thermoplastic	THW[d]	75°C 167°F 90°C 194°F	Dry and wet locations Special applications within electric discharge lighting equipment. Limited to 1000 open-circuit volts or less (size 14-8 only as permitted in 410.33)	Flame-retardant, moisture- and heat-resistant thermoplastic	14–10 8 6–2 1–4/0 213–500 501–1000 1001–2000	0.76 1.14 1.52 2.03 2.41 2.79 3.18	30 45 60 80 95 110 125	None

TABLE 19-1 Conductor Application and Insulations (*Continued*)

Trade Name	Type Letter	Maximum Operating Temperature	Application Provisions	Insulation	Thickness of insulation: AWG or kcmil	mm	mils	Outer Covering[a]
Moisture- and heat-resistant thermoplastic	THWN[d]	75°C 167°F	Dry and wet location	Flame-retardant, moisture- and heat-resistant thermoplastic	14–12 10 8–6 4–2 1–4/0 250–500 501–1000	0.38 0.51 0.76 1.02 1.27 1.52 1.78	15 20 30 40 50 60 70	Nylon jacket or equivalent
Moisture-resistant thermoplastic	TW	60°C 140°F	Dry and wet locations	Flame-retardant, moisture-resistant thermoplastic	14–10 8 6–2 1–4/0 213–500 501–1000 1001–2000	0.76 1.14 1.52 2.03 2.41 2.79 3.18	30 45 60 80 95 110 125	None
Underground feeder and branch-circuit cable—single conductor (for Type UF cable employing more than one conductor, see Article 340)	UF	60°C 140°F 75°C 167°F[7]	See Article 340	Moisture-resistant Moisture- and heat-resistant	14–10 8–2 1–4/0	1.52 2.03 2.41	60[f] 80[f] 95[f]	Integral with insulation
Underground service-entrance cable—single conductor (for Type USE cable employing more than one conductor, see Article 338)	USE[d]	75°C 167°F	See Article 338	Heat- and moisture-resistant	14–10 8–2 1–4/0 213–500 501–1000 1001–2000	1.14 1.52 2.03 2.41 2.79 3.18	5 60 80 95[h] 110 125	Moisture-resistant nonmetallic covering (See 338.2.)
Thermoset	XHH	90°C 194°F	Dry and damp location	Flame-retardant thermoset	14–10 8–2 1–4/0 213–500 501–1000 1001–2000	0.76 1.14 1.40 1.65 2.03 2.41	30 45 55 65 80 95	None
Moisture-resistant thermoset	XHHW[d]	90°C 194°F 167°F	Dry and damp location Wet location	Flame-retardant, moisture-resistant thermoset	14–10 8–2 1–4/0 213–500 501–1000 1001–2000	0.76 1.14 1.40 1.65 2.03 2.41	30 45 55 65 80 95	None
Moisture-resistant thermoset	XHHW-2	90°C 194°F	Dry and wet locations	Flame-retardant, moisture-resistant thermoset	14–10 8–2 1–4/0 213–500 501–1000 1001–2000	0.76 1.14 1.40 1.65 2.03 2.41	30 45 55 65 80 95	None
Modified ethylene tetrafluoro-ethylene	Z	90°C 194°F 150°C 302°F	Dry and damp locations Dry locations—special applications[b]	Modified ethylene tetrafluoro-ethylene	14–12 10 8–4 3–1 1/0–4/0	0.38 0.51 0.64 0.89 1.14	15 20 25 35 45	None

(*Continued*)

TABLE 19-1 Conductor Application and Insulations (*Continued*)

Trade Name	Type Letter	Maximum Operating Temperature	Application Provisions	Insulation	Thickness of insulation AWG or kcmil	mm	mils	Outer Covering[a]
Modified ethylene tetrafluoro-ethylene	ZW[d]	75°C 167°F	Wet locations	Modified ethylene tetrafluoro-ethylene	14–10	0.76	30	None
					8–2	1.14	45	
		90°C 194°F	Dry and damp locations					
		150°C 302°F	Dry locations—special applications[b]					

[a]Some insulations do not require an outer covering.
[b]Where design conditions require maximum conductor operating temperatures above 90°C (194°F).
[c]For signaling circuits permitting 300-V insulation.
[d]Listed wire types designated with the suffix "2," such as RHW-2, shall be permitted to be used at a continuous 90°C (194°F) operating temperature, wet or dry.
[e]Some rubber insulations do not require an outer covering.
[f]Includes integral jacket.
[g]For ampacity limitation, see 340.80.
[h]Insulation thickness shall be permitted to be 2.03 mm (80 mils) for listed Type USE conductors that have been subjected to special investigations. The nonmetallic covering over individual rubber-covered conductors of aluminum-sheathed cable and of lead-sheathed or multiconductor cable shall not be required to be flame retardant. For Type MC cable, see 330.104. For nonmetallic-sheathed cable, see Article 334, Part III. For Type UF cable, see Article 340, Part III.

Source: Reprinted with permission from NFPA 70-2005, the *National Electrical Code*®, © 2004 National Fire Protection Association, Quiney, Mass. 02269. This reprinted material is not the complete and official position of the National Fire Protection Association on the referenced subject, which is represented only by the standard in its entirety.

Cable-Assembly Systems. These are used extensively for concealed wiring not embedded in masonry or concrete. They may also be installed exposed in dry locations, and depending on the particular construction and ratings, in wet locations. Branch-circuit sizes are conventionally 600 V-rated. Cables rated for 5 through 15 kV are frequently used for primary distribution feeders in large commercial and industrial electrical systems.

In industrial plants and commercial utility areas, cable assemblies are often installed in expanded metal trays, ladder racks, or other approved cable-support systems.

Nonmetallic-sheathed cables are almost universally used in single family house wiring in the United States and in many multifamily occupancies. Armored cable is extensively used in commercial applications (see Fig. 19-2). Armored cable is used in extending branch circuits from outlet boxes on rigid conduit or EMT systems to lighting fixtures in suspended ceiling work.

Metal-clad type MC cable applies to constructions using interlocked armor, close fittings, or flexible corrugated tube over No. 18 copper, No. 12 aluminum, or larger conductors.

Two other metal-sheathed cables of special construction are recognized by the Code. Mineral-insulated metal-sheathed cable is sheathed with a continuous copper or steel outer covering, containing one or more conductors and insulated with highly compressed refractory mineral insulation. It is widely used in industrial power, control wiring and in either wet or dry locations.

MI must be terminated and connected by means of fittings designed and approved for the purpose.

Open wiring on knobs and cleats is rarely encountered in current work. Open feeders are still used in some industrial construction where low cost is a consideration, no safety hazard is involved, and appearance is unimportant (see Fig. 19-3).

Several cable assemblies have been developed for limited or particular uses, rather than for complete wiring systems for a building. The NEC should be consulted for specific requirements in each case.

Service-entrance (SE) cable is a form of armored or nonmetallic-sheathed cable specifically approved for service-entrance use. It is available in two types: SE, with a flame-retardant, moisture-resistant outer covering, and underground service-entrance cable suitable for direct burial in the ground.

TABLE 19-2A Allowable Ampacities of Single-Insulated Conductors, Rated 0 through 2000 V, 150°C through 250°C (302°F through 482°F), in Free Air, Based on Ambient Air Temperature of 40°C (104°F)

	Temperature Rating of Conductor (See Table 19-1)				
	150°C (302°F)	200°C (392°F)	250°C (482°F)	150°C (302°F)	
	Type Z	Types FEP, FEPB, PEA, SA	Types PFAH, TFE	Type Z	
Size AWG or kcmil	Copper		Nickel, or Nickel-coated Copper	Aluminum or Copper-clad Aluminum	Size AWG or kcmil
14	46	54	59	—	14
12	60	68	78	47	12
10	80	90	107	63	10
8	106	124	142	83	8
6	155	165	205	112	6
4	190	220	278	148	4
3	214	252	327	170	3
2	255	393	381	198	2
1	293	344	440	228	1
1/0	339	399	532	263	1/0
2/0	390	467	591	305	2/0
3/0	451	546	708	351	3/0
4/0	529	629	830	411	4/0
Correction Factors					
Ambient Temperature (°C)	For ambient temperatures other than 40°C (104°F), multiply the allowable ampacities shown above by the appropriate factor shown below				Ambient Temperature (°F)
41–50	0.95	0.97	0.98	0.95	105–122
51–60	0.90	0.94	0.95	0.90	123–140
61–70	0.85	0.90	0.93	0.85	141–158
71–80	0.80	0.87	0.90	0.80	159–176
81–90	0.74	0.83	0.87	0.74	177–194
91–100	0.67	0.79	0.85	0.67	195–212
101–120	0.52	0.71	0.79	0.52	213–248
121–140	0.30	0.61	0.72	0.30	249–284
141–160	—	0.50	0.65	—	285–320
161–180	—	0.35	0.58	—	321–356
181–200	—	—	0.49	—	357–392
201–225	—	—	0.35	—	393–437

Source: Table 19-2A to 19-2E reprinted with permission from NFPA 70-2005, the *National Electrical Code* © 2004, National Fire Protection Association, Quincy, Mass. 02269. This reprinted material is not the complete and official position of the National Fire Protection Association on the referenced subject, which is represented only by the standard in its entirety.

TABLE 19-2B Allowable Ampacities of Insulated Conductors Rated 0 through 2000 Volts, 60°C through 90°C (140°F through 194°F), Not More Than Three Current-Carrying Conductors in Raceway, Cable, or Earth (Directly Buried), Based on Ambient Temperature of 30°C (86°F)

	Temperature Rating of Conductor (See Table 19-1)						
	60°C (140°F)	75°C (167°F)	90°C (194°F)	60°C (140°F)	75°C (167°F)	90°C (194°F)	
	Types TW, UF	Types RHW, THHW, THW, THWN, XHHW, USE, ZW	Types TBS, SA, SIS, FEP, FEPB, MI, RHH, RHW-2, THHN, THHW, THW-2, THWN-2, USE-2, XHH, XHHW, XHHW-2, ZW-2	Types TW, UF	Types RHW, THHW, THW, THWN, XHHW, USE	Types TBS, SA, SIS, THHN, THHW, THW-2, THWN-2, RHH, RHW-2, USE-2, XHH, XHHW, XHHW-2, ZW-2	
Size AWG or kcmil	Copper			Aluminum or Copper-clad Aluminum			Size AWG or kcmil
18	—	—	14	—	—	—	—
16*	—	—	18	—	—	—	—
14*	20	20	25	—	—	—	—
12*	25	25	30	20	20	25	12*
10	30	35	40	25	30	35	10*
8	40	50	55	30	40	45	8
6	55	65	75	40	50	60	6
4	70	85	95	55	65	75	4
3	85	100	110	65	75	85	3
2	95	115	130	75	90	100	2
1	110	130	150	85	100	115	1
1/0	125	150	170	100	120	135	1/0
2/0	145	175	195	115	135	150	2/0
3/0	165	200	225	130	155	175	3/0
4/0	195	230	260	150	180	205	4/0
250	215	255	290	170	205	230	250
300	240	285	320	190	230	255	300
350	260	310	350	210	250	280	350
400	280	335	380	225	270	305	400
500	320	380	430	260	310	350	500
600	355	420	475	285	340	385	600
700	385	460	520	310	375	420	700
750	400	475	535	320	385	435	750
800	410	490	555	330	395	450	800
900	435	520	585	355	425	480	900
1000	455	545	615	375	445	500	1000
1250	495	590	665	405	485	545	1250
1500	520	625	705	435	520	585	1500
1750	545	650	735	455	545	615	1750
2000	560	665	750	470	560	630	2000

TABLE 19-2B Allowable Ampacities of Insulated Conductors Rated 0 through 2000 Volts, 60°C through 90°C (140°F through 194°F), Not More Than Three Current-Carrying Conductors in Raceway, Cable, or Earth (Directly Buried), Based on Ambient Temperature of 30°C (86°F) (*Continued*)

Correction Factors							
Ambient Temperature (°C)	For ambient temperatures other than 30°C (86°F), multiply the allowable ampacities shown above by the appropriate factor shown below						Ambient Temperature (°F)
21–25	1.08	1.05	1.04	1.08	1.05	1.04	70–77
26–30	1.00	1.00	1.00	1.00	1.00	1.00	78–86
31–35	0.91	0.94	0.96	0.91	0.94	0.96	87–95
36–40	0.82	0.88	0.91	0.82	0.88	0.91	96–104
41–45	0.71	0.82	0.87	0.71	0.82	0.87	105–113
46–50	0.58	0.75	0.82	0.58	0.75	0.82	114–122
51–55	0.41	0.67	0.76	0.41	0.67	0.76	123–131
56–60	—	0.58	0.71	—	0.58	0.71	132–140
61–70	—	0.33	0.58	—	0.33	0.58	141–158
71–80	—	—	0.41	—	—	0.41	159–176

*Unless specifically permitted, the overcurrent protection shall not exceed 15 A for 14 AWG, 20 A for 12 AWG, and 30 A for 10 AWG copper; or 15 A for 12 AWG and 25 A for 10 AWG aluminum and copper-clad aluminum after any correction factors for ambient temperature and number of conductors have been applied.

Extensions, Raceways, Conduits, Wireways, and Busways. *Nonmetallic surface extensions* are 2-wire assemblies limited to exposed work in office (or residence) occupancies, where additional outlets are to be installed in the same room with the outlet from which the extension originates. The location must be dry and not subject to corrosive vapors. The voltage should not exceed 150 V between conductors.

Underplaster extensions have been used as a concealed-wiring method to install additional outlets on an existing branch circuit. They were eliminated from the NEC in 1993 as a specific article since other articles addressed this method.

In general, the raceway systems were developed for special purposes and are of more commercial importance and find a more varied use than the special cable-assembly systems discussed earlier. This is particularly true of underfloor and cellular raceways for concealed work and of wireways and busways for exposed work. In cases where great flexibility in the use of electric power is of importance, the application of one of these special systems should be considered. In each case, the NEC should be consulted for specific installation rules.

Flexible-metal conduit, consisting of a flexible metallic tube roughly similar to the armor of armored cable, is used generally with rigid-conduit or electrical metallic tubing systems, to provide flexible connections at motor terminals, for instance, or in place of the rigid product where installations of the latter would be difficult owing to numerous bends, close working quarters, etc. The conductors are installed after the flexible conduit is in place.

Surface metal raceways (see Fig. 19-4) are flat, rectangular wireways used for exposed work in dry locations. They are frequently used to install additional outlets in a building already wired, where concealment of conductors is difficult, and are also used for special purposes, for example, installation of cove lighting and for show-window reflectors. Unless made of a metal at least 0.040 in thick, they are limited to use on circuits not exceeding 300 V.

Liquidtight flexible-metal conduit is, as the name suggests, a type of flexible-metal conduit having an outer jacket impervious to liquids and terminated in liquidtight fitting. It is most widely used for connecting motors to rigid-conduit systems or fixed-equipment enclosures.

TABLE 19-2C Allowable Ampacities of Single-Insulated Conductors Rated 0 through 2000 V in Free Air, Based on Ambient Air Temperature of 30°C (86°F)

	Temperature Rating of Conductor (See Table 19-1)						
	60°C (140°F)	75°C (167°F)	90°C (194°F)	60°C (140°F)	75°C (167°F)	90°C (194°F)	
	Types TW, UF	Types RHW, THHW, THW, THWN, XHHW, ZW	Types TBS, SA, SIS, FEP, FEPB, MI, RHH, RHW-2, THHN, THHW, THW-2, THWN-2, USE-2, XHH, XHHW, XHHW-2, ZW-2	Types TW, UF	Types RHW, THHW, THW, THWN, XHHW	Types TBS, SA, SIS, THHN, THHW, THW-2, THWN-2, RHH, RHW-2, USE-2, XHH, XHHW, XHHW-2, ZW-2	
Size AWG or kcmil	Copper			Aluminum or Copper-clad Aluminum			Size AWG or kcmil
18	—	—	18	—	—	—	—
16	—	—	24	—	—	—	—
14*	25	30	35	—	—	—	—
12*	30	35	40	25	30	35	12*
10*	40	50	55	35	40	40	10*
8	60	70	80	45	55	60	8
6	80	95	105	60	75	80	6
4	105	125	140	80	100	110	4
3	120	145	165	95	115	130	3
2	140	170	190	110	135	150	2
1	165	195	220	130	155	175	1
1/0	195	230	260	150	180	205	1/0
2/0	225	265	300	175	210	235	2/0
3/0	260	310	350	200	240	275	3/0
4/0	300	360	405	235	280	315	4/0
250	340	405	455	265	315	355	250
300	375	445	505	290	350	395	300
350	420	505	570	330	395	445	350
400	455	545	615	355	425	480	400
500	515	620	700	405	485	545	500
600	575	690	780	455	540	615	600
700	630	755	855	500	595	675	700
750	655	785	885	515	620	700	750
800	680	815	920	535	645	725	800
900	730	870	985	580	700	785	900
1000	780	935	1055	625	750	845	1000
1250	890	1065	1200	710	855	960	1250
1500	980	1175	1325	795	950	1075	1500
1750	1070	1280	1445	875	1050	1185	1750
2000	1155	1385	1560	960	1150	1335	2000

TABLE 19-2C Allowable Ampacities of Single-Insulated Conductors Rated 0 through 2000 V in Free Air, Based on Ambient Air Temperature of 30°C (86°F) (*Continued*)

Correction Factors							
Ambient Temperature (°C)	For ambient temperatures other than 30°C (86°F), multiply the allowable ampacities shown above by the appropriate factor shown below						Ambient Temperature (°F)
21–25	1.08	1.05	1.04	1.08	1.05	1.04	70–77
26–30	1.00	1.00	1.00	1.00	1.00	1.00	78–86
31–35	0.91	0.94	0.96	0.91	0.94	0.96	87–95
36–40	0.82	0.88	0.91	0.82	0.88	0.91	96–104
41–45	0.71	0.82	0.87	0.71	0.82	0.87	105–113
46–50	0.58	0.75	0.82	0.58	0.75	0.82	114–122
51–55	0.41	0.67	0.76	0.41	0.67	0.76	123–131
56–60	—	0.58	0.71	—	0.58	0.71	132–140
61–70	—	0.33	0.58	—	0.33	0.58	141–158
71–80	—	—	0.41	—	—	0.41	159–176

*Unless specifically permitted, the overcurrent protection shall not exceed 15 A for 14 AWG, 20 A for 12 AWG, and 30 A for 10 AWG copper; or 15 A for 12 AWG and 25 A for 10 AWG aluminum and copper-clad aluminum after any correction factors for ambient temperature and number of conductors have been applied.

Underfloor raceways (Fig. 19-5) are employed in buildings of fire-resistant construction to provide readily accessible raceways in the floor slab for light and power, telephone, and signal circuits. One, two, or three ducts are installed, depending on the desired uses. Junction boxes which mark each end of a run of raceway, and the tops of which are flush with the floor covering, make it possible to locate accurately the run of duct and, hence, to install additional outlets with the special tools provided by the manufacturer. Owing to its flexibility, this type of construction is particularly suitable for large office areas or where outlet locations are subject to change.

The *cellular-metal-floor raceway* involves a cellular-steel floor (Fig. 19-6*a*), which is a structural load-carrying element whose hollow cells form the wire raceway and a system of transverse headers, together with the necessary fittings and adapters. The headers are also wire raceways, providing electrical access from distribution points to any predetermined number of cells. The system can be designed to provide overall floor and ceiling electrical service for conductors not larger than No. 0 AWG, not only for light and power but also for telephone and signal circuits. The large internal-cell areas (normally on 6-in centers) afford adequate conductor space, while the complete floor and ceiling coverage provides for great flexibility in use during the building life, since access to headers and cells can be obtained at any time for additional outlets, new or rerouted circuits, etc.

Cellular-concrete-floor raceways are precast slabs with tubular "cells" designed to lineup in a continuous raceway. Cells terminate in metallic header ducts and other special fittings for connection to other parts of the electrical systems. Fittings approved for the purpose are inserted into the cell to provide for outlets (see Fig. 19-6*b*).

Structural raceways are formed-steel members which may be assembled to provide for the installation of electrical wires and cables. Such assemblies also provide for the installation of wiring devices in vertical members which may be concealed.

Wireways provide a convenient, exposed rectangular metal raceway or trough for no more than 30 current-carrying conductors or total conductor cross-sectional area not exceeding 20% of the interior cross-sectional area of the wireway. The product is available in several standard lengths, which are bolted together for continuous runs. Access at any point is through hinged covers and conduit knockouts. A complete array of fittings assures flexibility for various installation conditions.

Owing to their size, wireways can be used to advantage for large numbers of conductors, for a group of circuits leaving a branch-circuit panelboard or feeder distribution board.

TABLE 19-2D Allowable Ampacities of Insulated Conductors Rated 0 through 2000 V, 150°C through 250°C (302°F through 482°F). Not More Than Three Current-Carrying Conductors in Raceway or Cable, Based on Ambient Air Temperature of 40°C (104°F)

	Temperature Rating of Conductor (See Table 19-1)				
	150°C (302°F)	200°C (392°F)	250°C (482°F)	150°C (302°F)	
	Type Z	Types FEP, FEPB, PFA, SA	Types PFAH, TFE	Type Z	
Size AWG or kcmil	Copper		Nickel, or Nickel-coated Copper	Aluminum or Copper-clad Aluminum	Size AWG or kcmil
14	34	36	39	—	14
12	43	45	54	30	12
10	55	60	73	44	10
8	76	83	93	57	8
6	96	110	117	75	6
4	120	125	148	94	4
3	143	152	166	109	3
2	160	171	191	124	2
1	186	197	215	145	1
1/0	215	229	244	169	1/0
2/0	251	260	273	198	2/0
3/0	288	297	308	227	3/0
4/0	332	346	361	260	4/0
Correction Factors					
Ambient Temperature (°C)	For ambient temperatures other than 40°C (104°F), multiply the allowable ampacities shown above by the appropriate factor shown below				Ambient Temperature (°F)
41–50	0.95	0.97	0.98	0.95	105–122
51–60	0.90	0.94	0.95	0.90	123–140
61–70	0.85	0.90	0.93	0.85	141–158
71–80	0.80	0.87	0.90	0.80	159–176
81–90	0.74	0.83	0.87	0.74	177–194
91–100	0.67	0.79	0.85	0.67	195–212
101–120	0.52	0.71	0.79	0.52	213–248
121–140	0.30	0.61	0.72	0.30	249–284
141–160	—	0.50	0.65	—	285–320
161–180	—	0.35	0.58	—	321–356
181–200	—	—	0.49	—	357–392
201–225	—	—	0.35	—	393–437

TABLE 19-2E Ampacities of Not More Than Three Single Insulated Conductors, Rated 0 through 2000 V, Supported on a Messenger, Based on Ambient Air Temperature of 40°C (104°F)

	Temperature Rating of Conductor (See Table 19-1)				
	75°C (167°F)	90°C (194°F)	75°C (167°F)	90°C (194°F)	
	Types RHW, THHW, THW, THWN, XHHW, ZW	Types MI, THHN, THHW, THW-2, THWN-2, RHH, RHW-2, USE-2, XHHW, XHHW-2, ZW-2	Types RHW, THW, THWN, THHW, XHHW	Types THHN, THHW, RHH, XHHW, RHW-2 XHHW-2, THW-2 THWN-2, USE-2, ZW-2	
Size AWG or kcmil	Copper		Aluminum or Copper-clad Aluminum		Size AWG or kcmil
8	57	66	44	51	8
6	76	89	59	69	6
4	101	117	78	91	4
3	118	138	92	107	3
2	135	158	106	123	2
1	158	185	123	144	1
1/0	183	214	143	167	1/0
2/0	212	247	165	193	2/0
3/0	245	287	192	224	3/0
4/0	287	335	224	262	4/0
250	320	374	251	292	250
300	359	419	282	328	300
350	397	464	312	364	350
400	430	503	339	395	400
500	496	580	392	458	500
600	553	647	440	514	600
700	610	714	488	570	700
750	638	747	512	598	750
800	660	773	532	622	800
900	704	826	572	669	900
1000	748	879	612	716	1000
Correction Factors					
Ambient Temperature (°C)	For ambient temperatures other than 40°C (104°F), multiply the allowable ampacities shown above by the appropriate factor shown below				Ambient Temperature (°F)
21–25	1.20	1.14	1.20	1.14	70–77
26–30	1.13	1.10	1.13	1.10	79–86
31–35	1.07	1.05	1.07	1.05	88–95
36–40	1.00	1.00	1.00	1.00	97–104
41–45	0.93	0.95	0.93	0.95	106–113
46–50	0.85	0.89	0.85	0.89	115–122
51–55	0.76	0.84	0.76	0.84	124–131
56–60	0.65	0.77	0.65	0.77	133–140
61–70	0.38	0.63	0.38	0.63	142–158
71–80	—	0.45	—	0.45	160–176

TABLE 19-2F Ampacities of Bare or Covered Conductors in Free Air, Based on 40°C (104°F) Ambient, 80°C (176°F). Total Conductor Temperature, 610 mm/sec (2 ft/sec) Wind Velocity

Copper Conductors				AAC Aluminum Conductors			
Bare		Covered		Bare		Covered	
AWG or kcmil	Amperes	AWG or kcmil	Amperes	AWG or kcmil	Amperes	AWG or kcmil	Amperes
8	98	8	103	8	76	8	80
6	124	6	130	6	96	6	101
4	155	4	163	4	121	4	127
2	209	2	219	2	163	2	171
1/0	282	1/0	297	1/0	220	1/0	231
2/0	329	2/0	344	2/0	255	2/0	268
3/0	382	3/0	401	3/0	297	3/0	312
4/0	444	4/0	466	4/0	346	4/0	364
250	494	250	519	266.8	403	266.8	423
300	556	300	584	336.4	468	336.4	492
500	773	500	812	397.5	522	397.5	548
750	1000	750	1050	477.0	588	477.0	617
1000	1193	1000	1253	556.5	650	556.5	682
—	—	—	—	636.0	709	636.0	744
—	—	—	—	795.0	819	795.0	860
—	—	—	—	954.0	920	—	—
—	—	—	—	1033.5	968	1033.5	1017
—	—	—	—	1272	1103	1272	1201
—	—	—	—	1590	1267	1590	1381
—	—	—	—	2000	1454	2000	1527

Busways (Fig. 19-7) are one of the more important recent developments for exposed heavy-capacity feeder and circuit wiring in industrial plants because of their flexibility in use, which makes them readily adaptable to future needs and to changing conditions such as relocation or revamping of production lines. The initial investment can be confined to immediate requirements and additions made at anytime as requirements increase. The system consists essentially of interconnected prefabricated lengths or sections of steel or aluminum duct which enclose bus bars mounted on insulators. Regularly spaced openings in the sides of the duct permit plugging in branch-circuit control devices of the circuit-breaker, fuse, or fused-switch type, for convenient control of individual or group motor drives, lighting or heating circuits, etc. The ease of relocating both the duct and control devices makes its use advantageous for supplying power to machines on assembly lines, mass production manufacturing, and other applications where flexibility of electric supply is essential. Busways are available in capacities ranging from about 125 to about 3000 A, for 3-phase 3- or 4-wire systems.

The so-called trolley duct (Fig. 19-8) is a variation of the busway in which the metal duct and electrical buses (either single-phase or 3-phase) are so arranged that access is had to the buses at any point in the run. Current is collected from the buses by movable trolleys to which are wired portable or movable electrical devices. In industrial plants, the system is used to supply power to cranes and hoists, to portable tools on assembly lines and benches, etc. It has found some application in drafting rooms, stock departments, and similar locations, where ability to move lighting units quickly is of advantage.

Multioutlet assemblies are surface-mounted raceways of metal or plastic with plug receptacle outlets at spaced intervals or provisions for the insertion of receptacles as desired. Multioutlet assemblies are widely used where a number of cord-connected appliances must be served (as along the back of a workbench or laboratory table). They are also used to provide greater convenience for the attachment of portable cords. In this application, they are usually installed along the top of the baseboard (as around the perimeter of a private office).

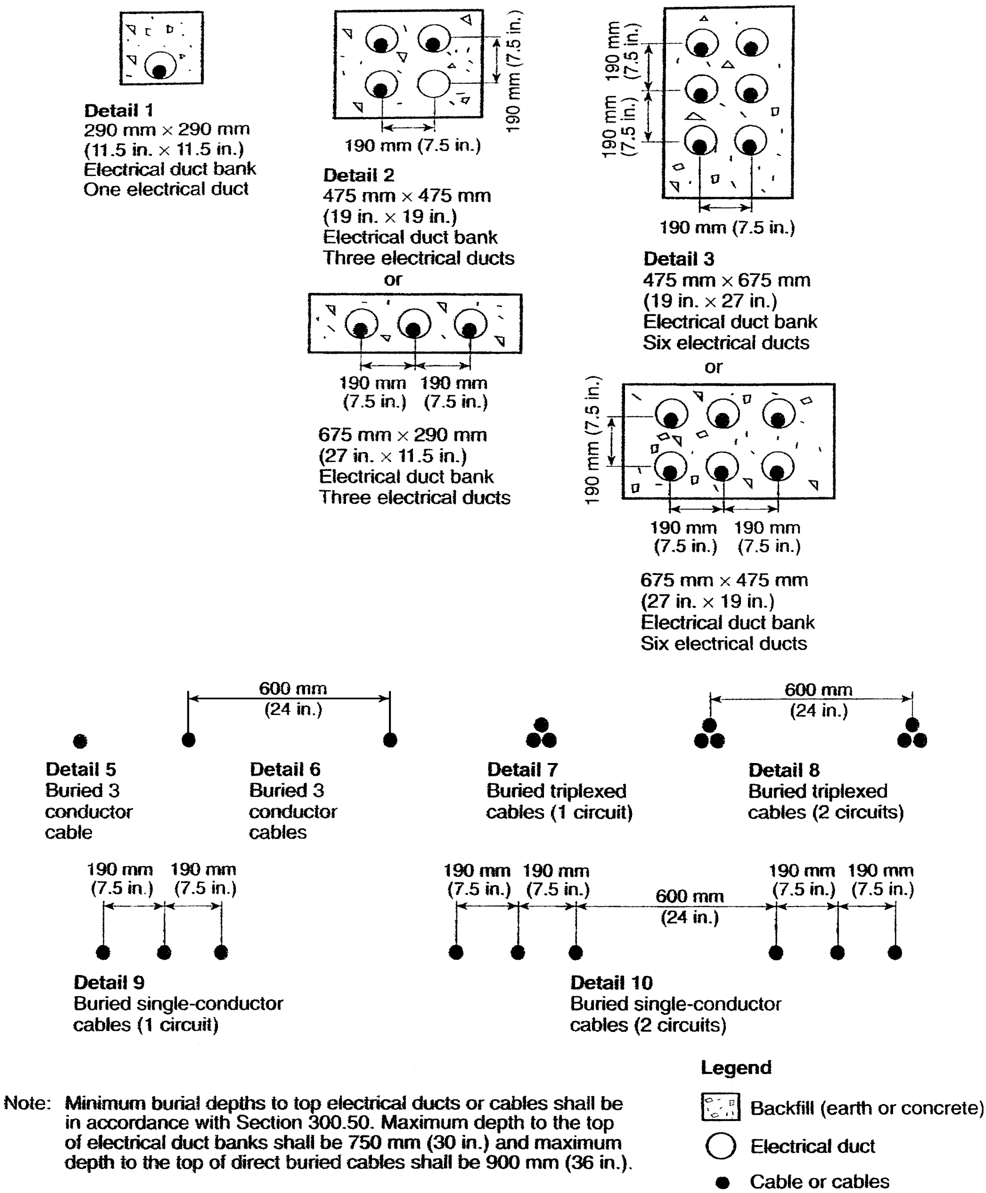


FIGURE 19-1 Configurations for buried systems using conductors in Tables 19-2A and 19-2C. (Reprinted with permission from NFPA 70-2005, *National Electrical Code*®, © 2004, National Fire Protection Association, Quincy, MA 02169. This reprinted material is not the complete and official position of the NFPA on the referenced subject, which is represented only by the standard in its entirety.

TABLE 19-3 Percent of Cross Section of Conduit and Tubbing for Conductors

Number of Conductors	All Conductors Types
1	53
2	31
Over 2	40

Based on common conditions of proper cabling and alignment of conductors where the length of the pull and the number of bends are within reasonable limits. It should be recognized that, for certain conditions, a larger size conduit or a lesser conduit fill should be considered.

When pulling three conductors or cables into a raceway, if the ratio of the raceway (inside diameter) to the conductor or cable (outside diameter) is between 2.8 and 3.2, jamming can occur. While jamming can occur when pulling four or more conductors or cables into a raceway, the probability is very low.

Conductors for Building Wiring. The various types of conductor available for interior wiring, together with their sizes, insulations, and uses, are indicated in Tables 19-1 and 19-2. Rubber and thermoplastic insulations are available in a number of compounds and constructions for resistance to heat, moisture, or other environmental conditions.

Other insulations used in building wiring include magnesium oxide, fluorinated ethylene propylene, silicone rubber, and the long-familiar varnished-cambric and asbestos constructions, although asbestos is no longer used except under special conditions.

Various connector types are shown in Fig. 19-9.

Dimensions of *insulated conductors* and fixture wires are given in Table 19-4A.

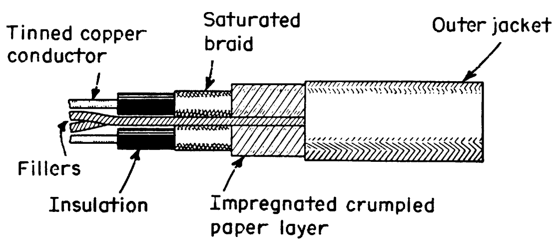


FIGURE 19-2 Nonmetallic sheathed cable.

Current-Carrying Capacity (Ampacity) and Other Properties. As the conductors of an electrical wiring system offer some resistance, a current-carrying conductor dissipates heat. Under practical conditions of installation and operation, the temperatures reached must not result in the destruction of the insulation or risk to surrounding material.

Tables of maximum allowable current-carrying capacity are given in the NEC. Allowable ampacities for insulated conductors are based on an allowable temperature rise above an ambient of 30°C (86°F) and 40°C (104°F). A list of temperature ratings for types of insulated conductors is given in Table 19-1.

Allowable ampacities for copper conductors and aluminum conductors in accordance with the temperature rating of the insulation are given for installation in conduit and for installation in free air in Tables 19-2A through 19-2E (see also Table 19-4B).

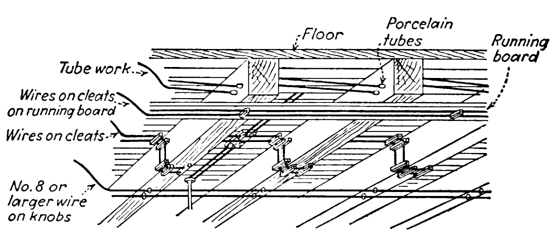


FIGURE 19-3 Methods of supporting open wiring.

To determine the permissible percent raceway fill for conductor combinations, see Table 19-3.*

Conductor and conduit diameters and areas are frequently necessary to calculate allowable fill. Nominal values for conductors are given in Table 19-4A and, for conduit and tubing. Tables 19-5A and 19-5B give resistance and reactance values of conductors in ohms per 304.8 m (1000 ft). Table 19-6 gives dimensions and area of conduit and tubing.

*Tables 19-3 through 19-7 are reprinted with permission from NFPA 70-2005, the *National Electrical Code*, © 2004, National Fire Protection Association, Quincy, Mass. 02269. This reprinted material is not the complete and official position of the National Fire Protection Association on the referenced subject which is represented only by the standard in its entirety.

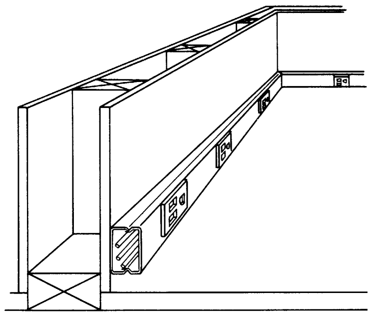


FIGURE 19-4 Typical surface raceway with plug receptacles.

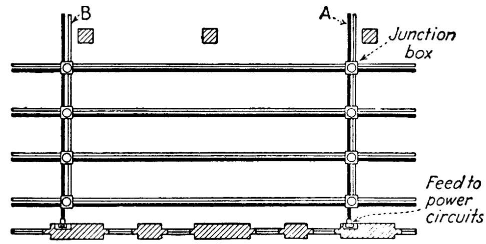


FIGURE 19-5 Layout of double underflow duct system: *A*—for power circuits; *B*—for signal and telephone circuits.

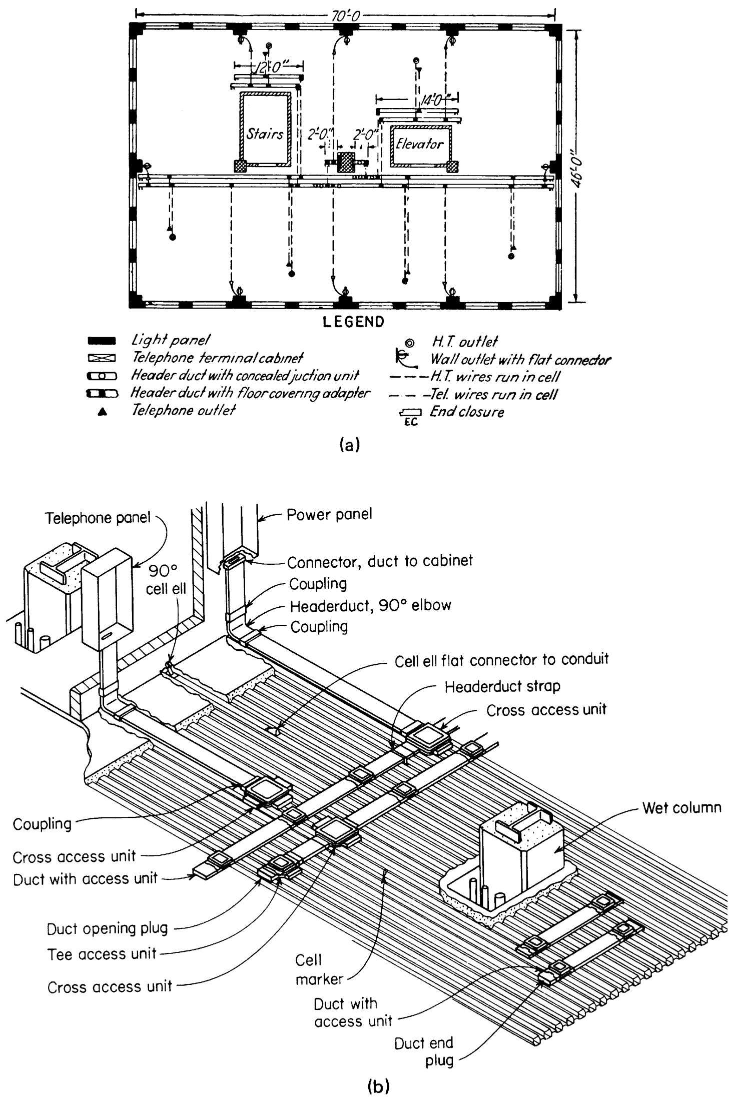


FIGURE 19-6 (*a*) Cellular-flow wiring layout; (*b*) floor ducts and access units.

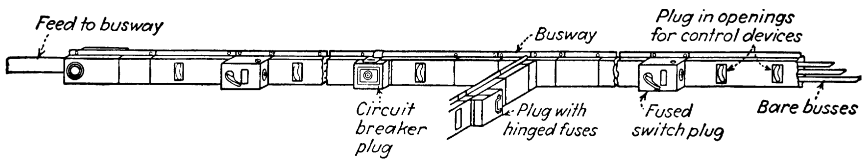


FIGURE 19-7 Units of busway distribution system.

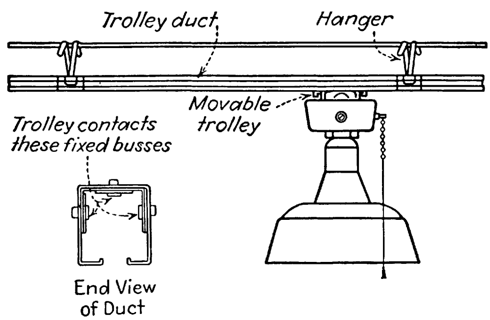


FIGURE 19-8 Trolley duct used for movable lighting fixture.

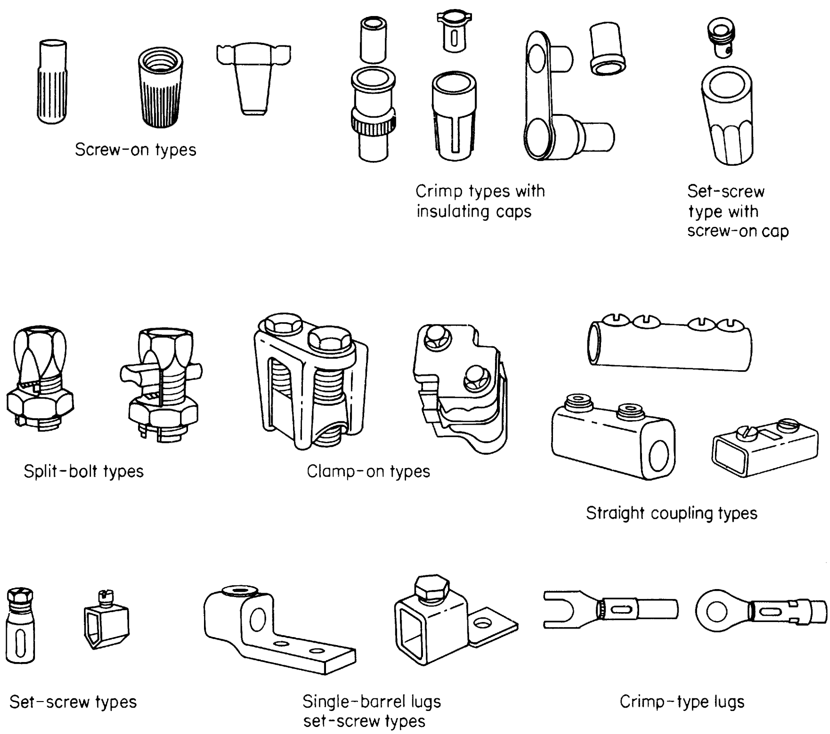


FIGURE 19-9 Types of wire connectors.

TABLE 19-4A Dimensions of Insulated Conductors and Fixture Wires

Type	Size (AWG or kcmil)	Approximate Diameter		Approximate Area	
		mm	in	mm²	in²
Type: FFH-2, RFH-1, RFH-2, RHH*, RHW*, RHW-2*, RHH, RHW, RHW-2, SF-1, SF-2, SFF-1, SFF-2, TF, TFF, THHW, THW, THW-2, TW, XF, XFF					
RFH-2, FFH-2	18	3.454	0.136	9.355	0.0145
	16	3.759	0.148	11.10	0.0172
RHH, RHW, RHW-2	14	4.902	0.193	18.90	0.0293
	12	5.385	0.212	22.77	0.0353
	10	5.994	0.236	28.19	0.0437
	8	8.280	0.326	53.87	0.0835
	6	9.246	0.364	67.16	0.1041
	4	10.46	0.412	86.00	0.1333
	3	11.18	0.440	98.13	0.1521
	2	11.99	0.472	112.9	0.1750
	1	14.78	0.582	171.6	0.2660
	1/0	15.80	0.622	196.1	0.3039
	2/0	16.97	0.668	226.1	0.3505
	3/0	18.29	0.720	262.7	0.4072
	4/0	19.76	0.778	306.7	0.4754
	250	22.73	0.895	405.9	0.6291
	300	24.13	0.950	457.3	0.7088
	350	25.43	1.001	507.7	0.7870
	400	26.62	1.048	556.5	0.8626
	500	28.78	1.133	650.5	1.0082
	600	31.57	1.243	782.9	1.2135
	700	33.38	1.314	874.9	1.3561
	750	34.24	1.348	920.8	1.4272
	800	35.05	1.380	965.0	1.4957
	900	36.68	1.444	1057	1.6377
	1000	38.15	1.502	1143	1.7719
	1250	43.92	1.729	1515	2.3479
	1500	47.04	1.852	1738	2.6938
	1750	49.94	1.966	1959	3.0357
	2000	52.63	2.072	2175	3.3719
SF-2, SFF-2	18	3.073	0.121	7.419	0.0115
	16	3.378	0.133	8.968	0.0139
	14	3.759	0.148	11.10	0.0172
SF-1, SFF-1	18	2.311	0.091	4.194	0.0065
RFH-1, XF, XFF	18	2.692	0.106	5.161	0.0080
TF, TFF, XF, XFF	16	2.997	0.118	7.032	0.0109
TW, XF, XFF, THHW, THW, THW-2	14	3.378	0.133	8.968	0.0139

(Continued)

TABLE 19-4A Dimensions of Insulated Conductors and Fixture Wires (*Continued*)

Type	Size (AWG or kcmil)	Approximate Diameter		Approximate Area	
		mm	in	mm^2	in^2
TW, THHW, THW, THW-2	12	3.861	0.152	11.68	0.0181
	10	4.470	0.176	15.68	0.0243
	8	5.994	0.236	28.19	0.0437
RHH*, RHW*, RHW-2*	14	4.140	0.163	13.48	0.0209
RHH*, RHW*, RHW-2*, XF, XFF	12	4.623	0.182	16.77	0.0260
Type: RHH*, RHW*, RHW-2*, THHN, THHW, THW, THW-2, TFN, TFFN, THWN, THWN-2, XF, XFF					
RHH*, RHW*, RHW-2*, XF, XFF	10	5.232	0.206	21.48	0.0333
RHH*, RHW*, RHW-2*	8	6.756	0.266	35.87	0.0556
TW, THW, THHW, THW-2, RHH*, RHW*, RHW-2*	6	7.722	0.304	46.84	0.0726
	4	8.941	0.352	62.77	0.0973
	3	9.652	0.380	73.16	0.1134
	2	10.46	0.412	86.00	0.1333
	1	12.50	0.492	122.6	0.1901
	1/0	13.51	0.532	143.4	0.2223
	2/0	14.68	0.578	169.3	0.2624
	3/0	16.00	0.630	201.1	0.3117
	4/0	17.48	0.688	239.9	0.3718
	250	19.43	0.765	296.5	0.4596
	300	20.83	0.820	340.7	0.5281
	350	22.12	0.871	384.4	0.5958
	400	23.32	0.918	427.0	0.6619
	500	25.48	1.003	509.7	0.7901
	600	28.27	1.113	627.7	0.9729
	700	30.07	1.184	710.3	1.1010
	750	30.94	1.218	751.7	1.1652
	800	31.75	1.250	791.7	1.2272
	900	33.38	1.314	874.9	1.3561
	1000	34.85	1.372	953.8	1.4784
	1250	39.09	1.539	1200	1.8602
	1500	42.21	1.662	1400	2.1695
	1750	45.11	1.776	1598	2.4773
	2000	47.80	1.882	1795	2.7818
TFN, TFFN	18	2.134	0.084	3.548	0.0055
	16	2.438	0.096	4.645	0.0072

TABLE 19-4A Dimensions of Insulated Conductors and Fixture Wires (*Continued*)

Type	Size (AWG or kcmil)	Approximate Diameter mm	Approximate Diameter in	Approximate Area mm^2	Approximate Area in^2
THHN, THWN, THWN-2	14	2.819	0.111	6.258	0.0097
	12	3.302	0.130	8.581	0.0133
	10	4.166	0.164	13.61	0.0211
	8	5.486	0.216	23.61	0.0366
	6	6.452	0.254	32.71	0.0507
	4	8.230	0.324	53.16	0.0824
	3	8.941	0.352	62.77	0.0973
	2	9.754	0.384	74.71	0.1158
	1	11.33	0.446	100.8	0.1562
	1/0	12.34	0.486	119.7	0.1855
	2/0	13.51	0.532	143.4	0.2223
	3/0	14.83	0.584	172.8	0.2679
	4/0	16.31	0.642	208.8	0.3237
	250	18.06	0.711	256.1	0.3970
	300	19.46	0.766	297.3	0.4608
Type: FEP, FEPB, PAF, PAFF, PF, PFA, PFAH, PFF, PGF, PGFF, PTF, PTFF, TFE, THHN, THWN, THWN-2, Z, ZF, ZFF					
THHN, THWN, THWN-2	350	20.75	0.817	338.2	0.5242
	400	21.95	0.864	378.3	0.5863
	500	24.10	0.949	456.3	0.7073
	600	26.70	1.051	559.7	0.8676
	700	28.50	1.122	637.9	0.9887
	750	29.36	1.156	677.2	1.0496
	800	30.18	1.188	715.2	1.1085
	900	31.80	1.252	794.3	1.2311
	1000	33.27	1.310	869.5	1.3478
PF, PGFF, PGF, PFF, PTF, PAF, PTFF, PAFF	18	2.184	0.086	3.742	0.0058
	16	2.489	0.098	4.839	0.0075
PF, PGFF, PGF, PFF, PTF, PAF, PTFF, PAFF, TFE, FEP, PFA, FEPB, PFAH	14	2.870	0.113	6.452	0.0100
TFE, FEP, PFA, FEPB, PFAH	12	3.353	0.132	8.839	0.0137
	10	3.962	0.156	12.32	0.0191
	8	5.232	0.206	21.48	0.0333
	6	6.198	0.244	30.19	0.0468
	4	7.417	0.292	43.23	0.0670
	3	8.128	0.320	51.87	0.0804
	2	8.941	0.352	62.77	0.0973
TFE, PFAH	1	10.72	0.422	90.26	0.1399
TFE, PFA, PFAH, Z	1/0	11.73	0.462	108.1	0.1676
	2/0	12.90	0.508	130.8	0.2027
	3/0	14.22	0.560	158.9	0.2463
	4/0	15.70	0.618	193.5	0.3000
ZF, ZFF	18	1.930	0.076	2.903	0.0045
	16	2.235	0.088	3.935	0.0061
Z, ZF, ZFF	14	2.616	0.103	5.355	0.0083

(*Continued*)

TABLE 19-4A Dimensions of Insulated Conductors and Fixture Wires (*Continued*)

Type	Size (AWG or kcmil)	Approximate Diameter		Approximate Area	
		mm	in	mm^2	in^2
Z	12	3.099	0.122	7.548	0.0117
	10	3.962	0.156	12.32	0.0191
	8	4.978	0.196	19.48	0.0302
	6	5.944	0.234	27.74	0.0430
	4	7.163	0.282	40.32	0.0625
	3	8.382	0.330	55.16	0.0855
	2	9.195	0.362	66.39	0.1029
	1	10.21	0.402	81.87	0.1269
Type: KF-1, KF-2, KFF-1, KFF-2, XHH, XHHW, XHHW-2, ZW					
XHHW, ZW, XHHW-2, XHH	14	3.378	0.133	8.968	0.0139
	12	3.861	0.152	11.68	0.0181
	10	4.470	0.176	15.68	0.0243
	8	5.994	0.236	28.19	0.0437
	6	6.960	0.274	38.06	0.0590
	4	8.179	0.322	52.52	0.0814
	3	8.890	0.350	62.06	0.0962
	2	9.703	0.382	73.94	0.1146
XHHW, XHHW-2, XHH	1	11.23	0.442	98.97	0.1534
	1/0	12.24	0.482	117.7	0.1825
	2/0	13.41	0.528	141.3	0.2190
	3/0	14.73	0.58	170.5	0.2642
	4/0	16.21	0.638	206.3	0.3197
	250	17.91	0.705	251.9	0.3904
	300	19.30	0.76	292.6	0.4536
	350	20.60	0.811	333.3	0.5166
	400	21.79	0.858	373.0	0.5782
	500	23.95	0.943	450.6	0.6984
	600	26.75	1.053	561.9	0.8709
	700	28.55	1.124	640.2	0.9923
	750	29.41	1.158	679.5	1.0532
	800	30.23	1.190	717.5	1.1122
	900	31.85	1.254	796.8	1.2351
	1000	33.32	1.312	872.2	1.3519
	1250	37.57	1.479	1108	1.7180
	1500	40.69	1.602	1300	2.0157
	1750	43.59	1.716	1492	2.3127
	2000	46.28	1.822	1682	2.6073
KF-2, KFF-2	18	1.600	0.063	2.000	0.0031
	16	1.905	0.075	2.839	0.0044
	14	2.286	0.090	4.129	0.0064
	12	2.769	0.109	6.000	0.0093
	10	3.378	0.133	8.968	0.0139
KF-1, KFF-1	18	1.448	0.057	1.677	0.0026
	16	1.753	0.069	2.387	0.0037
	14	2.134	0.084	3.548	0.0055
	12	2.616	0.103	5.355	0.0083
	10	3.226	0.127	8.194	0.0127

*Types RHH, RHW, and RHW-2 without outer covering

TABLE 19-4B Compact Aluminum Building Wire Nominal Dimensions* and Areas

Size (AWG or kcmil)	Bare Conductor Diameter		Types THW and THHW Approximate Diameter		Types THW and THHW Approximate Area		Type THHN Approximate Diameter		Type THHN Approximate Area		Type XHHW Approximate Diameter		Type XHHW Approximate Area		Size, (AWG or kcmil)
	mm	in	mm	in	mm^2	in^2	mm	in	mm^2	in^2	mm	in	mm^2	in^2	
8	3.404	0.134	6.477	0.255	32.90	0.0510	—	—	—	—	5.690	0.224	25.42	0.0394	8
6	4.293	0.169	7.366	0.290	42.58	0.0660	6.096	0.240	29.16	0.0452	6.604	0.260	34.19	0.0530	6
4	5.410	0.213	8.509	0.335	56.84	0.0881	7.747	0.305	47.10	0.0730	7.747	0.305	47.10	0.0730	4
2	6.807	0.268	9.906	0.390	77.03	0.1194	9.144	0.360	65.61	0.1017	9.144	0.360	65.61	0.1017	2
1	7.595	0.299	11.81	0.465	109.5	0.1698	10.54	0.415	87.23	0.1352	10.54	0.415	87.23	0.1352	1
1/0	8.534	0.336	12.70	0.500	126.6	0.1963	11.43	0.450	102.6	0.1590	11.43	0.450	102.6	0.1590	1/0
2/0	9.550	0.376	13.84	0.545	150.5	0.2332	12.57	0.495	124.1	0.1924	12.45	0.490	121.6	0.1885	2/0
3/0	10.74	0.423	14.99	0.590	176.3	0.2733	13.72	0.540	147.7	0.2290	13.72	0.540	147.7	0.2290	3/0
4/0	12.07	0.475	16.38	0.645	210.8	0.3267	15.11	0.595	179.4	0.2780	14.99	0.590	176.3	0.2733	4/0
250	13.21	0.520	18.42	0.725	266.3	0.4128	17.02	0.670	227.4	0.3525	16.76	0.660	220.7	0.3421	250
300	14.48	0.570	19.69	0.775	304.3	0.4717	18.29	0.720	262.6	0.4071	18.16	0.715	259.0	0.4015	300
350	15.65	0.616	20.83	0.820	340.7	0.5281	19.56	0.770	300.4	0.4656	19.30	0.760	292.6	0.4536	350
400	16.74	0.659	21.97	0.865	379.1	0.5876	20.70	0.815	336.5	0.5216	20.32	0.800	324.3	0.5026	400
500	18.69	0.736	23.88	0.940	447.7	0.6939	22.48	0.885	396.8	0.6151	22.35	0.880	392.4	0.6082	500
600	20.65	0.813	26.67	1.050	558.6	0.8659	25.02	0.985	491.6	0.7620	24.89	0.980	486.6	0.7542	600
700	22.28	0.877	28.19	1.110	624.3	0.9676	26.67	1.050	558.6	0.8659	26.67	1.050	558.6	0.8659	700
750	23.06	0.908	29.21	1.150	670.1	1.0386	27.31	1.075	585.5	0.9076	27.69	1.090	602.0	0.9331	750
900	23.37	0.999	31.09	1.224	759.1	1.1766	30.33	1.194	722.5	1.1196	29.69	1.169	692.3	1.0733	900
1000	26.92	1.060	32.64	1.285	836.6	1.2968	31.88	1.255	798.1	1.2370	31.24	1.230	766.6	1.1882	1000

*Dimensions are from industry sources.

TABLE 19-5A Conductor Properties

Size (AWG or kcmil)	Area		Conductors: Stranding			Conductors: Overall				Direct-Current Resistance at 75°C (167°F): Copper				Aluminum	
	Circular			Diameter		Diameter		Area		Uncoated		Coated			
	mm²	mils	Quantity	mm	in	mm	in	mm²	in²	ohm/km	ohm/kFT	ohm/km	ohm/kFT	ohm/km	ohm/kFT
18	0.823	1620	1	—	—	1.02	0.040	0.823	0.001	25.5	7.77	26.5	8.08	42.0	12.8
18	0.823	1620	7	0.39	0.015	1.16	0.046	1.06	0.002	26.1	7.95	27.7	8.45	42.8	13.1
16	1.31	2580	1	—	—	1.29	0.051	1.31	0.002	16.0	4.89	16.7	5.08	26.4	8.05
16	1.31	2580	7	0.49	0.019	1.46	0.058	1.68	0.003	16.4	4.99	17.3	5.29	26.9	8.21
14	2.08	4110	1	—	—	1.63	0.064	2.08	0.003	10.1	3.07	10.4	3.19	16.6	5.06
14	2.08	4110	7	0.62	0.024	1.85	0.073	2.68	0.004	10.3	3.14	10.7	3.26	16.9	5.17
12	3.31	6530	1	—	—	2.05	0.081	3.31	0.005	6.34	1.93	6.57	2.01	10.45	3.18
12	3.31	6530	7	0.78	0.030	2.32	0.092	4.25	0.006	6.50	1.98	6.73	2.05	10.69	3.25
10	5.261	10380	1	—	—	2.588	0.102	5.26	0.008	3.984	1.21	4.148	1.26	6.561	2.00
10	5.261	10380	7	0.98	0.038	2.95	0.116	6.76	0.011	4.070	1.24	4.226	1.29	6.679	2.04
8	8.367	16510	1	—	—	3.264	0.128	8.37	0.013	2.506	0.764	2.579	0.786	4.125	1.26
8	8.367	16510	7	1.23	0.049	3.71	0.146	10.76	0.017	2.551	0.778	2.653	0.809	4.204	1.28
6	13.30	26240	7	1.56	0.061	4.67	0.184	17.09	0.027	1.608	0.491	1.671	0.510	2.652	0.808
4	21.15	41740	7	1.96	0.077	5.89	0.232	27.19	0.042	1.010	0.308	1.053	0.321	1.666	0.508
3	26.67	52620	7	2.20	0.087	6.60	0.260	34.28	0.053	0.802	0.245	0.833	0.254	1.320	0.403
2	33.62	66360	7	2.47	0.097	7.42	0.292	43.23	0.067	0.634	0.194	0.661	0.201	1.045	0.319
1	42.41	83690	19	1.69	0.066	8.43	0.332	55.80	0.087	0.505	0.154	0.524	0.160	0.829	0.253
1/0	53.49	105600	19	1.89	0.074	9.45	0.372	70.41	0.109	0.399	0.122	0.415	0.127	0.660	0.201
2/0	67.43	133100	19	2.13	0.084	10.62	0.418	88.74	0.137	0.3170	0.0967	0.329	0.101	0.523	0.159
3/0	85.01	167800	19	2.39	0.094	11.94	0.470	111.9	0.173	0.2512	0.0766	0.2610	0.0797	0.413	0.126
4/0	107.2	211600	19	2.68	0.106	13.41	0.528	141.1	0.219	0.1996	0.0608	0.2050	0.0626	0.328	0.100
250	127	—	37	2.09	0.082	14.61	0.575	168	0.260	0.1687	0.0515	0.1753	0.0535	0.2778	0.0847
300	152	—	37	2.29	0.090	16.00	0.630	201	0.312	0.1409	0.0429	0.1463	0.0446	0.2318	0.0707
350	177	—	37	2.47	0.097	17.30	0.681	235	0.364	0.1205	0.0367	0.1252	0.0382	0.1984	0.0605
400	203	—	37	2.64	0.104	18.49	0.728	268	0.416	0.1053	0.0321	0.1084	0.0331	0.1737	0.0529
500	253	—	37	2.95	0.116	20.65	0.813	336	0.519	0.0845	0.0258	0.0869	0.0265	0.1391	0.0424
600	304	—	61	2.52	0.099	22.68	0.893	404	0.626	0.0704	0.0214	0.0732	0.0223	0.1159	0.0353

700	355	—	61	2.72	0.107	24.49	0.964	471	0.730	0.0603	0.0184	0.0622	0.0189	0.0994	0.0303
750	380	—	61	2.82	0.111	25.35	0.998	505	0.782	0.0563	0.0171	0.0579	0.0176	0.0927	0.0282
800	405	—	61	2.91	0.114	26.16	1.030	538	0.834	0.0528	0.0161	0.0544	0.0166	0.0868	0.0265
900	456	—	61	3.09	0.122	27.79	1.094	606	0.940	0.0470	0.0143	0.0481	0.0147	0.0770	0.0235
1000	507	—	61	3.25	0.128	29.26	1.152	673	1.042	0.0423	0.0129	0.0434	0.0132	0.0695	0.0212
1250	633	—	91	2.98	0.117	32.74	1.289	842	1.305	0.0338	0.0103	0.0347	0.0106	0.0554	0.0169
1500	760	—	91	3.26	0.128	35.86	1.412	1011	1.566	0.02814	0.00858	0.02814	0.00883	0.0464	0.0141
1750	887	—	127	2.98	0.117	38.76	1.526	1180	1.829	0.02410	0.00735	0.02410	0.00756	0.0397	0.0121
2000	1013	—	127	3.19	0.126	41.45	1.632	1349	2.092	0.02109	0.00643	0.02109	0.00662	0.0348	0.0106

Notes:

1. These resistance values are valid only for the parameters as given. Using conductors having coated strands, different stranding type, and, especially, other temperatures changes the resistance.
2. Formula for temperature change: $R_2 = R_1 [1 + \alpha (T_2 - 75)]$ where $\alpha_{cu} = 0.00323$, $\alpha_{AL} = 0.00330$ at 75°C.
3. Conductors with compact and compressed stranding have about 9% and 3% respectively, smaller bare conductor diameters than those shown. See Table 5A for actual compact cable dimensions.
4. The IACS conductivities used: bare copper = 100%, aluminum = 61%.
5. Class B stranding is listed as well as solid for some sizes. Its overall diameter and area is that of its circumscribing circle.

FPN: The construction information is per NEMA WC8-1992 or ANSI/UL 1581-1998. The resistance is calculated per National Bureau of Standards Handbook 100, dated 1966, and Handbook 109, dated 1972.

TABLE 19-5B Alternating-Current Resistance and Reactance for 600-V Cables, 3-Phase, 60 Hz, 75°C (167°F)—Three Single Conductors in Conduit

Size (AWG or kcmil)	Ohms to Neutral per Kilometer / Ohms to Neutral per 1000 Feet																Size (AWG or kcmil)
	X_L (Reactance) for All Wires		Alternating-Current Resistance for Uncoated Copper Wires			Alternating-Current Resistance for Aluminum Wires			Effective Z at 0.85 PF for Uncoated Copper Wires			Effective Z at 0.85 PF for Aluminum Wires					
	PVC, Aluminum Conduits	Steel Conduit	PVC Conduit	Aluminum Conduit	Steel Conduit	PVC Conduit	Aluminum Conduit	Steel Conduit	PVC Conduit	Aluminum Conduit	Steel Conduit	PVC Conduit	Aluminum Conduit	Steel Conduit			
14	0.190 0.058	0.240 0.073	10.2 3.1	10.2 3.1	10.2 3.1	— —	— —	— —	8.9 2.7	8.9 2.7	8.9 2.7	— —	— —	— —	14		
12	0.177 0.054	0.223 0.068	6.6 2.0	6.6 2.0	6.6 2.0	10.5 3.2	10.5 3.2	10.5 3.2	5.6 1.7	5.6 1.7	5.6 1.7	9.2 2.8	9.2 2.8	9.2 2.8	12		
10	0.164 0.050	0.207 0.063	3.9 1.2	3.9 1.2	3.9 1.2	6.6 2.0	6.6 2.0	6.6 2.0	3.6 1.1	3.6 1.1	3.6 1.1	5.9 1.8	5.9 1.8	5.9 1.8	10		
8	0.171 0.052	0.213 0.065	2.56 0.78	2.56 0.78	2.56 0.78	4.3 1.3	4.3 1.3	4.3 1.3	2.26 0.69	2.26 0.69	2.30 0.70	3.6 1.1	3.6 1.1	3.6 1.1	8		
6	0.167 0.051	0.210 0.064	1.61 0.49	1.61 0.49	1.61 0.49	2.66 0.81	2.66 0.81	2.66 0.81	1.44 0.44	1.48 0.45	1.48 0.45	2.33 0.71	2.36 0.72	2.36 0.72	6		
4	0.157 0.048	0.197 0.060	1.02 0.31	1.02 0.31	1.02 0.31	1.67 0.51	1.67 0.51	1.67 0.51	0.95 0.29	0.95 0.29	0.98 0.30	1.51 0.46	1.51 0.46	1.51 0.46	4		
3	0.154 0.047	0.194 0.059	0.82 0.25	0.82 0.25	0.82 0.25	1.31 0.40	1.35 0.41	1.31 0.40	0.75 0.23	0.79 0.24	0.79 0.24	1.21 0.37	1.21 0.37	1.21 0.37	3		
2	0.148 0.045	0.187 0.057	0.62 0.19	0.66 0.20	0.66 0.20	1.05 0.32	1.05 0.32	1.05 0.32	0.62 0.19	0.62 0.19	0.66 0.20	0.98 0.30	0.98 0.30	0.98 0.30	2		
1	0.151 0.046	0.187 0.057	0.49 0.15	0.52 0.16	0.52 0.16	0.82 0.25	0.85 0.26	0.82 0.25	0.52 0.16	0.52 0.16	0.52 0.16	0.79 0.24	0.79 0.24	0.82 0.25	1		
1/0	0.144 0.044	0.180 0.055	0.39 0.12	0.43 0.13	0.39 0.12	0.66 0.20	0.69 0.21	0.66 0.20	0.43 0.13	0.43 0.13	0.43 0.13	0.62 0.19	0.66 0.20	0.66 0.20	1/0		
2/0	0.141 0.043	0.177 0.054	0.33 0.10	0.33 0.10	0.33 0.10	0.52 0.16	0.52 0.16	0.52 0.16	0.36 0.11	0.36 0.11	0.36 0.11	0.52 0.16	0.52 0.16	0.52 0.16	2/0		

3/0	0.138 0.042	0.171 0.052	0.253 0.077	0.269 0.082	0.259 0.079	0.43 0.13	0.43 0.13	0.43 0.13	0.289 0.088	0.302 0.092	0.308 0.094	0.43 0.13	0.43 0.13	0.46 0.14	3/0
4/0	0.135 0.041	0.167 0.051	0.203 0.062	0.220 0.067	0.207 0.063	0.33 0.10	0.36 0.11	0.33 0.10	0.243 0.074	0.256 0.078	0.262 0.080	0.36 0.11	0.36 0.11	0.36 0.11	4/0
250	0.135 0.041	0.171 0.052	0.171 0.052	0.187 0.057	0.177 0.054	0.279 0.085	0.295 0.090	0.282 0.086	0.217 0.066	0.230 0.070	0.240 0.073	0.308 0.094	0.322 0.098	0.33 0.10	250
300	0.135 0.041	0.167 0.051	0.144 0.044	0.161 0.049	0.148 0.045	0.233 0.071	0.249 0.076	0.236 0.072	0.194 0.059	0.207 0.063	0.213 0.065	0.269 0.082	0.282 0.086	0.289 0.088	300
350	0.131 0.040	0.164 0.050	0.125 0.038	0.141 0.043	0.128 0.039	0.200 0.061	0.217 0.066	0.207 0.063	0.174 0.053	0.190 0.058	0.197 0.060	0.240 0.073	0.253 0.077	0.262 0.080	350
400	0.131 0.040	0.161 0.049	0.108 0.033	0.125 0.038	0.115 0.035	0.177 0.054	0.194 0.059	0.180 0.055	0.161 0.049	0.174 0.053	0.184 0.056	0.217 0.066	0.233 0.071	0.240 0.073	400
500	0.128 0.039	0.157 0.048	0.089 0.027	0.105 0.032	0.095 0.029	0.141 0.043	0.157 0.048	0.148 0.045	0.141 0.043	0.157 0.048	0.164 0.050	0.187 0.057	0.200 0.061	0.210 0.064	500
600	0.128 0.039	0.157 0.048	0.075 0.023	0.092 0.028	0.082 0.025	0.118 0.036	0.135 0.041	0.125 0.038	0.131 0.040	0.144 0.044	0.154 0.047	0.167 0.051	0.180 0.055	0.190 0.058	600
750	0.125 0.038	0.157 0.048	0.062 0.019	0.079 0.024	0.069 0.021	0.095 0.029	0.112 0.034	0.102 0.031	0.118 0.036	0.131 0.040	0.141 0.043	0.148 0.045	0.161 0.049	0.171 0.052	750
1000	0.121 0.037	0.151 0.046	0.049 0.015	0.062 0.019	0.059 0.018	0.075 0.023	0.089 0.027	0.082 0.025	0.105 0.032	0.118 0.036	0.131 0.040	0.128 0.039	0.138 0.042	0.151 0.046	1000

Notes:

1. These values are based on the following constants: UL-Type RHH wires with Class B stranding, in cradled configuration. Wire conductivities are 100% IACS copper and 61% IACS aluminum, and aluminum conduit is 45% IACS. Capacitive reactance is ignored, since it is negligible at these voltages. These resistance values are valid only at 75°C (167°F) and for the parameters as given, but are representative for 600-V wire types operating at 60 Hz.

2. *Effective Z* is defined as $R\cos(\theta) + X\sin(\theta)$, where θ is the power factor angle of the circuit. Multiplying current by effective impedance gives a good approximation for line-to-neutral voltage drop. Effective impedance values shown in this table are valid only at 0.85 power factor. For another circuit power factor (*PF*), effective impedance (*Ze*) can be calculated from *R* and X_L values given in this table as follows: $Ze = R \times PF + X_L \sin[\arccos(PF)]$.

TABLE 19-6 Dimensions and Percent Area of Conduit and Tubing (Areas of Conduit or Tubing for the Combinations of Wires Permitted in Table 1, Chapter 9)

Article 358—Electrical Metallic Tubing (EMT)

Metric Designator	Trade Size	Nominal Internal Diameter		Total Area 100%		60%		1 Wire 53%		2 Wires 31%		Over 2 Wires 40%	
		mm	in	mm^2	in^2	mm^2	in^2	mm^2	in^2	mm^2	in^2	mm^2	in^2
16	1/2	15.8	0.622	196	0.304	118	0.182	104	0.161	61	0.094	78	0.122
21	3/4	20.9	0.824	343	0.533	206	0.320	182	0.283	106	0.165	137	0.213
27	1	26.6	1.049	556	0.864	333	0.519	295	0.458	172	0.268	222	0.346
35	1 1/4	35.1	1.380	968	1.496	581	0.897	513	0.793	300	0.464	387	0.598
41	1 1/2	40.9	1.610	1314	2.036	788	1.221	696	1.079	407	0.631	526	0.814
53	2	52.5	2.067	2165	3.356	1299	2.013	1147	1.778	671	1.040	866	1.342
63	2 1/2	69.4	2.731	3783	5.858	2270	3.515	2005	3.105	1173	1.816	1513	2.343
78	3	85.2	3.356	5701	8.846	3421	5.307	3022	4.688	1767	2.742	2280	3.538
91	3 1/2	97.4	3.834	7451	11.545	4471	6.927	3949	6.119	2310	3.579	2980	4.618
103	4	110.1	4.334	9521	14.753	5712	8.852	5046	7.819	2951	4.573	3808	5.901

Article 362—Electrical Nonmetallic Tubing (ENT)

Metric Designator	Trade Size	Nominal Internal Diameter		Total Area 100%		60%		1 Wire 53%		2 Wires 31%		Over 2 Wires 40%	
		mm	in	mm^2	in^2	mm^2	in^2	mm^2	in^2	mm^2	in^2	mm^2	in^2
16	1/2	14.2	0.560	158	0.246	95	0.148	84	1.131	49	0.076	63	0.099
21	3/4	19.3	0.760	293	0.454	176	0.272	155	0.240	91	0.141	117	0.181
27	1	25.4	1.000	507	0.785	304	0.471	269	0.416	157	0.243	203	0.314
35	1 1/4	34.0	1.340	908	1.410	545	0.846	481	0.747	281	0.437	363	0.564
41	1 1/2	39.9	1.570	1250	1.936	750	1.162	663	1.026	388	0.600	500	0.774
53		51.3	2.020	2067	3.205	1240	1.923	1095	1.699	641	0.993	827	1.282
63	2 1/2	—	—	—	—	—	—	—	—	—	—	—	—
78	3	—	—	—	—	—	—	—	—	—	—	—	—
91	3 1/4	—	—	—	—	—	—	—	—	—	—	—	—

Article 348—Flexible Metal Conduit (FMC)

Metric Designator	Trade Size	Nominal Internal Diameter		Total Area 100%		60%		1 Wire 53%		2 Wires 31%		Over 2 Wires 40%	
		mm	in	mm^2	in^2	mm^2	in^2	mm^2	in^2	mm^2	in^2	mm^2	in^2
12	3/8	9.7	0.384	74	0.116	44	0.069	39	0.061	23	0.036	30	0.046
16	1/2	16.1	0.635	204	0.317	122	0.190	108	0.168	63	0.098	81	0.127
21	3/4	20.9	0.824	343	0.533	206	0.320	182	0.283	106	0.165	137	0.213
27	1	25.9	1.020	527	0.817	316	0.490	279	0.433	163	0.253	211	0.327
35	1 1/4	32.4	1.275	824	1.277	495	0.766	437	0.677	256	0.396	330	0.511
41	1 1/2	39.1	1.538	1201	1.858	720	1.115	636	0.985	372	0.576	480	0.743
53	2	51.8	2.040	2107	3.269	1264	1.961	1117	1.732	653	1.013	843	1.307
63	2 1/2	63.5	2.500	3167	4.909	1900	2.945	1678	2.602	982	1.522	1267	1.963
78	3	76.2	3.000	4560	7.069	2736	4.241	2417	3.746	1414	2.191	1824	2.827
91	3 1/2	88.9	3.500	6207	9.621	3724	5.773	3290	5.099	1924	2.983	2483	3.848
103	4	101.6	4.000	8107	12.566	4864	7.540	4297	6.660	2513	3.896	3243	5.027

TABLE 19-6 Dimensions and Percent Area of Conduit and Tubing (Areas of Conduit or Tubing for the Combinations of Wires Permitted in Table 1, Chapter 9) (*Continued*)

Articles 352 and 353—Rigid PVC Conduit (RNC), Schedule 40, and HDPE Conduit

Metric Designator	Trade Size	Nominal Internal Diameter mm	Nominal Internal Diameter in	Total Area 100% mm²	Total Area 100% in²	60% mm²	60% in²	1 Wire 53% mm²	1 Wire 53% in²	2 Wires 31% mm²	2 Wires 31% in²	Over 2 Wires 40% mm²	Over 2 Wires 40% in²
12	3/8	—	—	—	—	—	—	—	—	—	—	—	—
16	1/2	15.3	0.602	184	0.285	110	0.171	97	0.151	57	0.088	74	0.114
21	3/4	20.4	0.804	327	0.508	196	0.305	173	0.269	101	0.157	131	0.203
27	1	26.1	1.029	535	0.832	321	0.499	284	0.441	166	0.258	214	0.333
35	1 1/4	34.5	1.360	935	1.453	561	0.872	495	0.770	290	0.450	374	0.581
41	1 1/2	40.4	1.590	1282	1.986	769	1.191	679	1.052	397	0.616	513	0.794
53	2	52.0	2.047	2124	3.291	1274	1.975	1126	1.744	658	1.020	849	1.316
63	2 1/2	62.1	2.445	3029	4.695	1817	2.817	1605	2.488	939	1.455	1212	1.878
78	3	77.3	3.042	4693	7.268	2816	4.361	2487	3.852	1455	2.253	1877	2.907
91	3 1/2	89.4	3.521	6277	9.737	3766	5.842	3327	5.161	1946	3.018	2511	3.895
103	4	101.5	3.998	8091	12.554	4855	7.532	4288	6.654	2508	3.892	3237	5.022
129	5	127.4	5.016	12748	19.761	7649	11.856	6756	10.473	3952	6.126	5099	7.904
155	6	153.2	6.031	18433	28.567	11060	17.140	9770	15.141	5714	8.856	7373	11.427

Article 352—Type A, Rigid PVC Conduit (RNC)

Metric Designator	Trade Size	Nominal Internal Diameter mm	Nominal Internal Diameter in	Total Area 100% mm²	Total Area 100% in²	60% mm²	60% in²	1 Wire 53% mm²	1 Wire 53% in²	2 Wires 31% mm²	2 Wires 31% in²	Over 2 Wires 40% mm²	Over 2 Wires 40% in²
16	1/2	17.8	0.700	249	0.385	149	0.231	132	0.204	77	0.119	100	0.154
21	3/4	23.1	0.910	419	0.650	251	0.390	222	0.345	130	0.202	168	0.260
27	1	29.8	1.175	697	1.084	418	0.651	370	0.575	216	0.336	279	0.434
35	1 1/4	38.1	1.500	1140	1.767	684	1.060	604	0.937	353	0.548	456	0.707
41	1 1/2	43.7	1.720	1500	2.324	900	1.394	795	1.231	465	0.720	600	0.929
53	2	54.7	2.155	2350	3.647	1410	2.188	1245	1.933	728	1.131	940	1.459
63	2 1/2	66.9	2.635	3515	5.453	2109	3.272	1863	2.890	1090	1.690	1406	2.181
78	3	82.0	3.230	5281	8.194	3169	4.916	2799	4.343	1637	2.540	2112	3.278
91	3 1/2	93.7	3.690	6896	10.694	4137	6.416	3655	5.668	2138	3.315	2758	4.278
103	4	106.2	4.180	8858	13.723	5315	8.234	4695	7.273	2746	4.254	3543	5.489
129	5	—	—	—	—	—	—	—	—	—	—	—	—
155	6	—	—	—	—	—	—	—	—	—	—	—	—

Article 352—Type EB, PVC Conduit (RNC)

Metric Designator	Trade Size	Nominal Internal Diameter mm	Nominal Internal Diameter in	Total Area 100% mm²	Total Area 100% in²	60% mm²	60% in²	1 Wire 53% mm²	1 Wire 53% in²	2 Wires 31% mm²	2 Wires 31% in²	Over 2 Wires 40% mm²	Over 2 Wires 40% in²
16	1/2	—	—	—	—	—	—	—	—	—	—	—	—
21	3/4	—	—	—	—	—	—	—	—	—	—	—	—
27	1	—	—	—	—	—	—	—	—	—	—	—	—
35	1 1/4	—	—	—	—	—	—	—	—	—	—	—	—
41	1 1/2	—	—	—	—	—	—	—	—	—	—	—	—
53	2	56.4	2.221	2498	3.874	1499	2.325	1324	2.053	774	1.201	999	1.550
63	2 1/2	—	—	—	—	—	—	—	—	—	—	—	—

(*Continued*)

TABLE 19-6 Dimensions and Percent Area of Conduit and Tubing (Areas of Conduit or Tubing for the Combinations of Wires Permitted in Table 1, Chapter 9) (*Continued*)

Article 352—Type EB, PVC Conduit (RNC)

Metric Designator	Trade Size	Nominal Internal Diameter mm	in	Total Area 100% mm²	in²	60% mm²	in²	1 Wire 53% mm²	in²	2 Wires 31% mm²	in²	Over 2 Wires 40% mm²	in²
78	3	84.6	3.330	5621	8.709	3373	5.226	2979	4.616	1743	2.700	2248	3.484
91	3½	96.6	3.804	7329	11.365	4397	6.819	3884	6.023	2272	3.523	2932	4.546
103	4	108.9	4.289	9314	14.448	5589	8.669	4937	7.657	2887	4.479	3726	5.779
129	5	135.0	5.316	14314	22.195	8588	13.317	7586	11.763	4437	6.881	5726	8.878
155	6	160.9	6.336	20333	31.530	12200	18.918	10776	16.711	6303	9.774	8133	12.612

Article 350—Liquidtight Flexible Metal Conduit (LFMC)

Metric Designator	Trade Size	Nominal Internal Diameter mm	in	Total Area 100% mm²	in²	60% mm²	in²	1 Wire 53% mm²	in²	2 Wires 31% mm²	in²	Over 2 Wires 40% mm²	in²
12	⅜	12.5	0.494	123	0.192	74	0.115	65	0.102	38	0.059	49	0.0?
16	½	16.1	0.632	204	0.314	122	0.188	108	0.166	63	0.097	81	0.0
21	¾	21.1	0.830	350	0.541	210	0.325	185	0.287	108	0.168	140	0.0
27	1	26.8	1.054	564	0.873	338	0.524	299	0.462	175	0.270	226	0.0
35	1¼	35.4	1.395	984	1.528	591	0.917	522	0.810	305	0.474	394	0.0
41	1½	40.3	1.588	1276	1.981	765	1.188	676	1.050	395	0.614	510	0.0
53	2	51.6	2.033	2091	3.246	1255	1.948	1108	1.720	648	1.006	836	1.29
63	2½	63.3	2.493	3147	4.881	1888	2.929	1668	2.587	976	1.513	1259	1.95
78	3	78.4	3.085	4827	7.475	2896	4.485	2559	3.962	1497	2.317	1931	2.99
91	3½	89.4	3.520	6277	9.731	3766	5.839	3327	5.158	1946	3.017	2511	3.89
103	4	102.1	4.020	8187	12.692	4912	7.615	4339	6.727	2538	3.935	3275	5.07
129	5	—	—	—	—	—	—	—	—	—	—	—	—
155	6	—	—	—	—	—	—	—	—	—	—	—	—

Article 344—Rigid Metal Conduit (RMC)

Metric Designator	Trade Size	Nominal Internal Diameter mm	in	Total Area 100% mm²	in²	60% mm²	in²	1 Wire 53% mm²	in²	2 Wires 31% mm²	in²	Over 2 Wires 40% mm²	in²
12	⅜	—	—	—	—	—	—	—	—	—	—	—	—
16	½	16.1	0.632	204	0.314	122	0.188	108	0.166	63	0.097	81	0.125
21	¾	21.2	0.836	353	0.549	212	0.329	187	0.291	109	0.170	141	0.220
27	1	27.0	1.063	573	0.887	344	0.532	303	0.470	177	0.275	229	0.355
35	1¼	35.4	1.394	984	1.526	591	0.916	522	0.809	305	0.473	394	0.610
41	1½	41.2	1.624	1333	2.071	800	1.243	707	1.098	413	0.642	533	0.829
53	2	52.9	2.083	2198	3.408	1319	2.045	1165	1.806	681	1.056	879	1.363
63	2½	63.2	2.489	3137	4.866	1882	2.919	1663	2.579	972	1.508	1255	1.946
78	3	78.5	3.090	4840	7.499	2904	4.499	2565	3.974	1500	2.325	1936	3.000
91	3½	90.7	3.570	6461	10.010	3877	6.006	3424	5.305	2003	3.103	2584	4.004
103	4	102.9	4.050	8316	12.882	4990	7.729	4408	6.828	2578	3.994	3326	5.153
129	5	128.9	5.073	13050	20.212	7830	12.127	6916	10.713	4045	6.266	5220	8.085
155	6	154.8	6.093	18821	29.158	11292	17.495	9975	15.454	5834	9.039	7528	11.663

TABLE 19-6 Dimensions and Percent Area of Conduit and Tubing (Areas of Conduit or Tubing for the Combinations of Wires Permitted in Table 1, Chapter 9) (*Continued*)

Article 352—Rigid PVC Conduit (RNC), Schedule 80

Metric Designator	Trade Size	Nominal Internal Diameter		Total Area 100%		60%		1 Wire 53%		2 Wires 31%		Over 2 Wires 40%	
		mm	in	mm²	in²	mm²	in²	mm²	in²	mm²	in²	mm²	in²
12	3/8	—	—	—	—	—	—	—	—	—	—	—	—
16	1/2	13.4	0.526	141	0.217	85	0.130	75	0.115	44	0.067	56	0.087
21	3/4	18.3	0.722	263	0.409	158	0.246	139	0.217	82	0.127	105	0.164
27	1	23.8	0.936	445	0.688	267	0.413	236	0.365	138	0.213	178	0.275
35	1 1/4	31.9	1.255	799	1.237	480	0.742	424	0.656	248	0.383	320	0.495
41	1 1/2	37.5	1.476	1104	1.711	663	1.027	585	0.907	342	0.530	442	0.684
53	2	48.6	1.913	1855	2.874	1113	1.725	983	1.523	575	0.891	742	1.150
63	2 1/2	58.2	2.290	2660	4.119	1596	2.471	1410	2.183	825	1.277	1064	1.647
78	3	72.7	2.864	4151	6.442	2491	3.865	2200	3.414	1287	1.997	1660	2.577
91	3 1/2	84.5	3.326	5608	8.688	3365	5.213	2972	4.605	1738	2.693	2243	3.475
103	4	96.2	3.786	7268	11.258	4361	6.755	3852	5.967	2253	3.490	2907	4.503
129	5	121.1	4.768	11518	17.855	6911	10.713	6105	9.463	3571	5.535	4607	7.142
155	6	145.0	5.709	16513	25.598	9908	15.359	8752	13.567	5119	7.935	6605	10.239

Article 342—Intermediate Metal Conduit (IMC)

Metric Designator	Trade Size	Nominal Internal Diameter		Total Area 100%		60%		1 Wire 53%		2 Wires 31%		Over 2 Wires 40%	
		mm	in	mm²	in²	mm²	in²	mm²	in²	mm²	in²	mm²	in²
12	3/8	—	—	—	—	—	—	—	—	—	—	—	—
16	1/2	16.8	0.660	222	0.342	133	0.205	117	0.181	69	0.106	89	0.137
21	3/4	21.9	0.864	377	0.586	226	0.352	200	0.311	117	0.182	151	0.235
27	1	28.1	1.105	620	0.959	372	0.575	329	0.508	192	0.297	248	0.384
35	1 1/4	36.8	1.448	1064	1.647	638	0.988	564	0.873	330	0.510	425	0.659
41	1 1/2	42.7	1.683	1432	2.225	859	1.335	759	1.179	444	0.690	573	0.890
53	2	54.6	2.150	2341	3.630	1405	2.178	1241	1.924	726	1.125	937	1.452
63	2 1/2	64.9	2.557	3308	5.135	1985	3.081	1753	2.722	1026	1.592	1323	2.054
78	3	80.7	3.176	5115	7.922	3069	4.753	2711	4.199	1586	2.456	2046	3.169
91	3 1/2	93.2	3.671	6822	10.584	4093	6.351	3616	5.610	2115	3.281	2729	4.234
103	4	105.4	4.166	8725	13.631	5235	8.179	4624	7.224	2705	4.226	3490	5.452

Article 356—Liquidtight Flexible Nonmetallic Conduit (LFNC-B*)

Metric Designator	Trade Size	Nominal Internal Diameter		Total Area 100%		60%		1 Wire 53%		2 Wires 31%		Over 2 Wires 40%	
		mm	in	mm²	in²	mm²	in²	mm²	in²	mm²	in²	mm²	in²
12	3/8	12.5	0.494	123	0.192	74	0.115	65	0.102	38	0.059	49	0.077
16	1/2	16.1	0.632	204	0.314	122	0.188	108	0.166	63	0.097	81	0.125
21	3/4	21.1	0.830	350	0.541	210	0.325	185	0.287	108	0.168	140	0.216
27	1	26.8	1.054	564	0.873	338	0.524	299	0.462	175	0.270	226	0.349
35	1 1/4	35.4	1.395	984	1.528	591	0.917	522	0.810	305	0.474	394	0.611
41	1 1/2	40.3	1.588	1276	1.981	765	1.188	676	1.050	395	0.614	510	0.792
53	2	51.6	2.033	2091	3.246	1255	1.948	1108	1.720	648	1.006	836	1.298

*Corresponds to 356.2(2)

(*Continued*)

TABLE 19-6 Dimensions and Percent Area of Conduit and Tubing (Areas of Conduit or Tubing for the Combinations of Wires Permitted in Table 1, Chapter 9) (*Continued*)

Article 356—Liquidtight Flexible Nonmetallic Conduit (LFNC-A*)

Metric Designator	Trade Size	Nominal Internal Diameter mm	Nominal Internal Diameter in	Total Area 100% mm²	Total Area 100% in²	60% mm²	60% in²	1 Wire 53% mm²	1 Wire 53% in²	2 Wires 31% mm²	2 Wires 31% in²	Over 2 Wires 40% mm²	Over 2 Wires 40% in²
12	3/8	12.6	0.495	125	0.192	75	0.115	66	0.102	39	0.060	50	0.077
16	1/2	16.0	0.630	201	0.312	121	0.187	107	0.165	62	0.097	80	0.125
21	3/4	21.0	0.825	346	0.535	208	0.321	184	0.283	107	0.166	139	0.214
27	1	26.5	1.043	552	0.854	331	0.513	292	0.453	171	0.265	221	0.342
35	1 1/4	35.1	1.383	968	1.502	581	0.901	513	0.796	300	0.466	387	0.601
41	1 1/2	40.7	1.603	1301	2.018	781	1.211	690	1.070	403	0.626	520	0.807
53	2	52.4	2.063	2157	3.343	1294	2.006	1143	1.772	669	1.036	863	1.337

*Corresponds to 356.2(1)

The number of conductors in one conduit or tubing is another important factor. The number of conductors of a certain size that may be installed in a given-sized conduit or EMT is limited to provide for ready installation and withdrawal without injury to conductor or insulating covering. Table 19-3 gives these values for commonly used conductors.

In considering the ampacity of conductors and conduit fill, it is important to note that derating of ampacity applies for increased ambient temperatures and for more than three conductors, excluding neutrals, in a conduit.

Flexible cords and fixture wire in some cases may be as small as No. 18 AWG; hence these are exceptions to the general rule that no conductor smaller than No. 14 should be used in light and power wiring. Owing to the heat generated in the lamp, heat-resistant wiring is required in fixtures.

19.4 TYPES OF CIRCUIT

Services and Feeders. No limit is placed on the electrical capacity of service conductors and service protection employed in bringing the electric supply into a building, since only one supply should be introduced whenever possible. Near the point of entrance of the supply, the heavy-service conductors are tapped by feeders, which conduct the electricity to panelboards at various load centers in the building, where the final branch circuits which supply individual lighting, heating, and power outlets originate (see Fig. 19-10). No limits are placed on the electrical capacity of feeders, but for practical purposes they are limited in size by the difficulty of handling large conductors and raceways in restricted building spaces, by voltage drop, and by economic considerations.

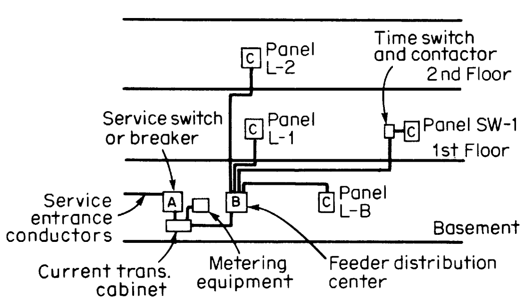


FIGURE 19-10 Riser diagram showing location of *A*— service, *B*—feeder, and *C*—branch circuit overcurrent protective devices.

TABLE 19-7 Summary of Branch-Circuit Requirements

Circuit Rating	15 A	20 A	30 A	40A	50 A
Conductors (min. size):					
Circuit wires*	14	12	10	8	6
Taps	14	14	14	12	12
Fixture wires and cords—see	240.5				
Overcurrent Protection	15A	20 A	30 A	40A	50 A
Outlet device:					
Lampholders permitted	Any type	Any type	Heavy duty	Heavy duty	Heavy duty
Receptacle rating†	15 max. A	15 or 20 A	30 A	40 or 50 A	50 A
Maximum Load	15A	20 A	30 A	40A	50 A
Permissible load	See 210.23(A)	See 210.23(A)	See 210.23(B)	See 210.23(C)	See 210.23(C)

*These gauges are for copper conductors.
†For receptacle rating of cord-connected electric-discharge luminaires (lighting fixtures), see NEC Section 410.30(C).

Each lighting fixture, motor, heating device, or other item of utilization equipment must be supplied by one of the types of branch circuit.

Branch Circuits for Grouped Loads. The uses and limitations of the common types of branch circuit are outlined in Table 19-7. It will be noted that lighting branch circuits may carry loads as high as 50 A, although fluorescent lighting is limited to use on circuits of 15- or 20-A rating. Such circuits are extensively employed in commercial and industrial occupancies. Branch circuits supplying convenience outlets for general use in other than manufacturing areas are usually limited to a maximum of 20 A, as the type of outlet required for heavier-capacity circuits usually will not accommodate the connection plug found on portable cords or lamps, motor-driven office machinery, etc.

Individual Branch Circuits. Any individual piece of equipment (except motors) may also be connected to a branch circuit meeting the following requirements: conductors must be large enough for the individual load supplied. Overcurrent protection must not exceed the capacity of the conductors or 150% of the rating of the individual load, if the single load device is a non-motor-operated appliance rated at 13.3 A or more. Only a single outlet or piece of equipment may be supplied.

Motor Branch Circuits. Owing to the peculiar conditions obtained during the starting period of a motor, and because it may be subjected to severe overloads at frequent intervals, motors, except for very small sizes, are connected to branch circuits of a somewhat different design from that previously discussed.

19.5 OVERCURRENT PROTECTION

Molded-Case Circuit Breakers. A circuit breaker is defined by the NEC as "a device designed to open and close a circuit by nonautomatic means and to open the circuit automatically on a predetermined overcurrent without damage to itself when properly applied within its ratings." In addition, the purpose of a circuit breaker (Fig. 19-11), as stated in the NEC, is to open the circuit if the current reaches a value that will cause an excessive or dangerous temperature in a conductor or its insulation.

FIGURE 19-11 Typical molded-case circuit breakers.

Underwriters' Laboratories (UL) Standard UL489, *Molded Case Circuit Breakers and Circuit Breaker Enclosures*, further specifies test and construction requirements for molded-case circuit breakers specifically intended to provide service entrance and feeder and branch circuit protection in accordance with the NEC. These UL requirements cover molded-case circuit breakers rated through 600 V. Underwriter's Laboratories specifies the applicable standards to which the circuit breaker is designed, whereas the NEC discusses proper applications of circuit breakers.

Ratings. The ratings which apply to circuit breakers and their actual assigned numerical values reflect mechanical, electrical, and thermal capabilities of those circuit breakers that comply with industry standards published by the National Electrical Manufacturers Association and UL. These ratings, which appear on the breaker, include the following:

1. *Voltage*. Circuit breakers are designed and marked with the maximum voltage at which they can be applied. They can be used on any system where the voltage is lower than this breaker rating. This includes both ac and dc voltage systems.
2. *Frequency*. Circuit breakers are normally suitable for use on 50- and 60-Hz electrical distribution systems. Rerating of the circuit breaker may be required at other frequencies.
3. *Continuous current*. Standard molded-case circuit breakers are calibrated to carry 100% of their rated current in free air at 40°C ambient. In accordance with the NEC, when installed in their enclosures, these breakers should not be continuously loaded over 80% of their current rating. However, there are certain molded-case circuit breakers specifically approved for 100% continuous rating available, and these are so marked.
4. *Current interrupting*. The current-interrupting rating is expressed in rms symmetrical amperes. It may vary with the applied voltage and is the maximum current the breaker can be expected to safely interrupt. These current-interrupting ratings may vary from 1500 through 200,000 A, depending upon applied voltage. Underwriter's Laboratory requires all circuit breakers to be so marked with their proper ratings.

Thermal-Magnetic Circuit Breakers. Thermal-magnetic circuit breakers provide two forms of overcurrent protection. The first is overload protection, which is achieved by a bimetal providing an inverse time-current response. The second is overcurrent protection, which is achieved magnetically.

WIRING DESIGN FOR COMMERCIAL AND INDUSTRIAL BUILDINGS **19-37**

Overload tripping is obtained through deflection of the bimetal, which is heated by the load current. During an overload condition, the bimetal deflects, causing the breaker to trip or open mechanically. The larger the overload, the faster the tripping of the circuit breaker; the smaller the overload, the longer it takes the circuit breaker to trip. This is commonly called an *inverse-time principle*.

Overcurrent (short circuit) protection is obtained through electromagnetic tripping action without any intentional time delay. An overcurrent condition must be interrupted rapidly (usually less than 20 ms) in order to protect downstream equipment. During overcurrent conditions, an armature is moved by electromagnetic force and initiates tripping action.

All circuit breakers are rated for continuous current. This continuous-current rating is dependent upon the maximum current the load is expected to draw for 3 h or more continuously. Ratings may be fixed or adjustable, depending on the circuit breaker. Some constructions may require replacing all or part of the trip unit to change the continuous current rating. Overcurrent trip characteristics are a function, multiple, or percentage of the continuous-current rating.

FIGURE 19-12 Current-limiting circuit breaker.

Current-Limiting Circuit Breakers. The NEC defines a current-limiting overcurrent protective device as "a device which, when interrupting current in its current limiting range, will reduce the current flow in the faulted circuit to a magnitude substantially less than that obtainable in the same circuit if the device were replaced with a solid conductor having comparable impedance."

Additionally, UL has the following current-limiting circuit-breaker definition: "a circuit breaker that does not employ a fusible element and that when operating within its current-limiting range, limits the let-through I^2 to a value less than the I^2 of a 1/2 cycle wave of the symmetrical prospective current."

Current-limiting circuit breakers (Fig. 19-12) not only provide high interrupting capability but also limit let-through current and energy to downstream devices. I^2t is an expression related to the energy resulting from current flow. Specific manufacturer's literature should be consulted for information on the current-limiting and energy-limiting characteristics of their particular circuit breakers.

Current-limiting circuit breakers provide the system designer with a means of reducing fault current and energy levels at downstream system components while still retaining the advantages, such as common trip and reusability, of circuit breaker protection. Current-limiting breakers can be reset and service restored in the same manner as conventional thermal-magnetic circuit breakers. There is nothing to replace even after clearing maximum fault currents. Figure 19-13 illustrates the current waveform resulting from current-limiting operation. At current levels below the threshold of current limitation, conventional non-current-limiting operation is the same as described for thermal-magnetic circuit breakers.

Solid-State Trip Circuit Breakers. Solid-state trip circuit breakers utilize solid-state electronic components to measure and time the output from current transformers. The solid-state circuitry then initiates tripping action based on predetermined settings. These components are housed in the trip unit section of the breaker.

The basic function of the trip unit is to provide long-time current-time delay and instantaneous current response characteristics necessary for proper circuit protection. Combinations of these characteristics provide time delay to override transient overload, delayed tripping for sustained overloads, and instantaneous tripping for high-level short circuits.

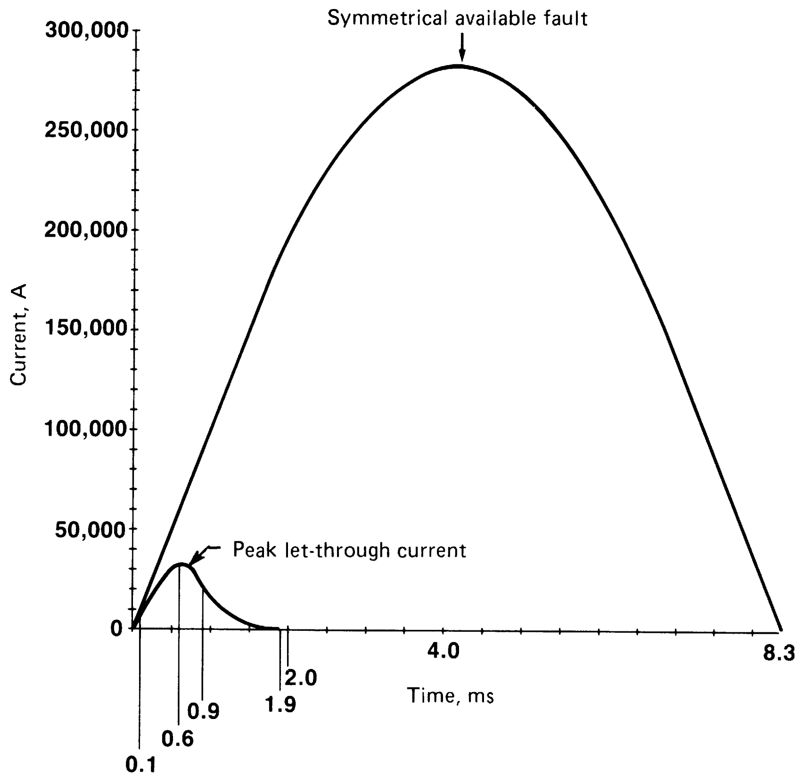


FIGURE 19-13 Time-current curve for a current-limiting breaker.

Solid-state trip units may also be equipped with short-time pickup and delay functions. These characteristics are useful for providing coordination with downstream or upstream circuit breakers and/or fuses. The resulting combination of long- and short-time delay characteristics provide delayed tripping for all levels of overcurrent below the instantaneous response. This provides time for downstream breakers to operate and clear the fault.

Solid-state trip units may also provide time-current characteristics for equipment ground fault protection in accordance with the NEC Sec. 230-95 requirements. The maximum current setting is 1200 A per NEC requirements, and time-delay settings are usually adjustable to permit coordination with other protective devices.

Time-Current Characteristic Curves. Circuit breaker time-current characteristic curves are a function of the type of trip function and its associated settings. A typical time-current characteristic curve is shown in Fig. 19-14 for a 400-A frame thermal-magnetic molded-case circuit breaker.

This particular curve applies for the continuous-current rating of 400 A and indicates total operational time from fault current initiation to clearing. This is a typical characteristic curve for a molded-case circuit breaker where overload sensing is achieved through a thermal element (a bimetal) and where instantaneous tripping is achieved magnetically.

In circuit-breaker frame sizes larger than 100 A, instantaneous operation is adjustable within approximately 5 to 10 times the continuous ampere rating.

Molded-case circuit breakers with more complex time-current characteristics are usually equipped with solid-state trip units. The continuous current rating may be adjustable using the long-time pickup function. In addition, long-time delay and short-time pickup and delay functions may be adjustable. The instantaneous response may also be adjustable.

Circuit breakers with solid-state trip units may also include integral equipment ground fault protection with current pickup and time-delay adjustments. Figure 19-15 illustrates a combined time-current characteristic curve for a solid-state electronic trip circuit breaker with all of these adjustments.

Further refinement of the solid-state trip unit short-time delay characteristics for overcurrent protection or the ground fault time-delay characteristics may be obtained by making part of the response curve a function of the product of time and the square of current (I^2t) illustrated by the sloping portions in Fig. 19-16. This sloping response curve often more readily coordinates with upstream or downstream thermal-magnetic circuit breakers and/or fuses.

It should be noted that solid-state trip units are available with varying degrees of complexity, and not all functions are provided for most applications. The specific manufacturer's literature should be consulted for guidance on individual units.

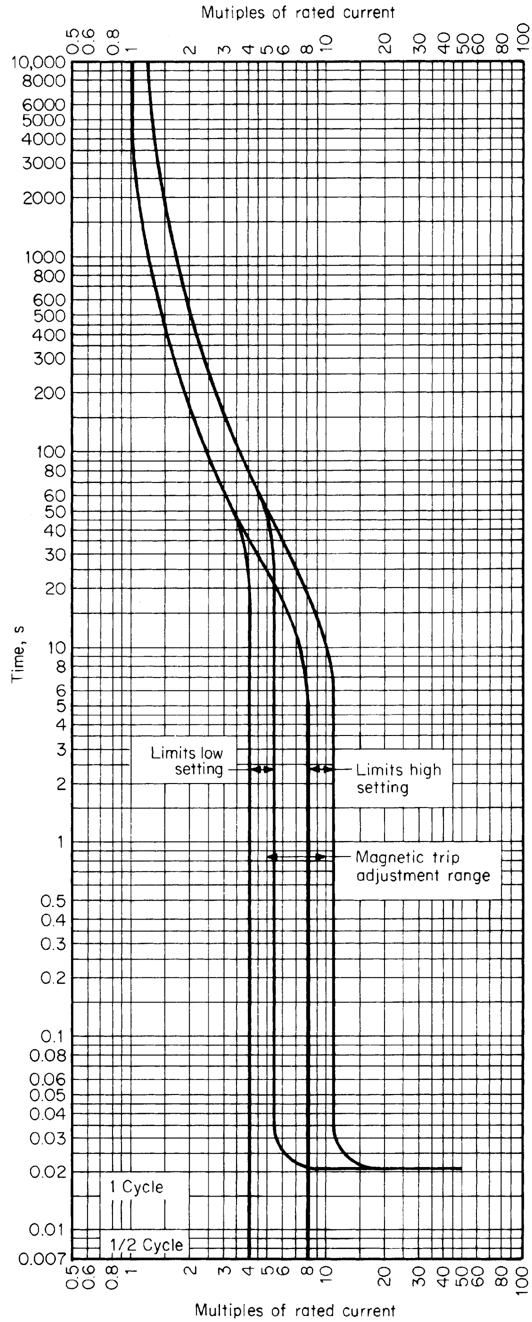


FIGURE 19-14 Characteristic tripping curve for molded-case breaker.

Series-Connected Ratings. *Series connection* of molded-case circuit breakers is a viable protection scheme, provided testing has verified performance. Underwriter's Laboratory presently recognizes series-connected short-circuit ratings and specifies test procedures to verify performance. Series-connected ratings must be based on tests and are only valid for specific circuit breaker types listed in the UL test reports. Individual manufacturer's series-connected ratings may be found in the *UL Recognized Component Directory*. Fuse-breaker coordinated combinations are also tested by UL and are applicable within their established guidelines.

Coordination. When protection is being considered, the performance of a circuit breaker with respect to the connected

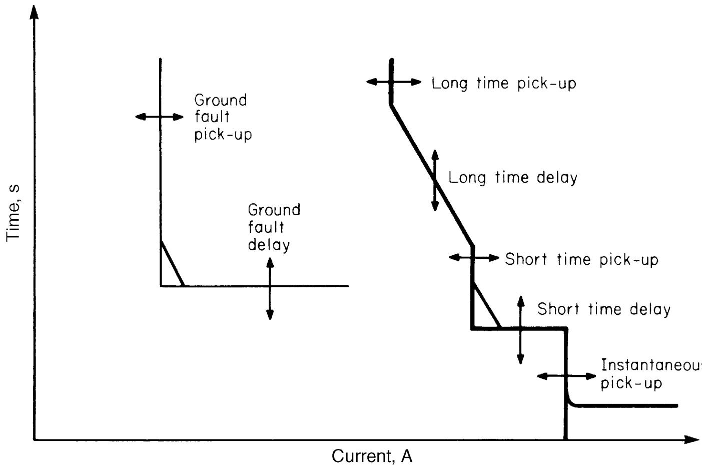


FIGURE 19-15 Time-current curve for solid-state electronic breaker.

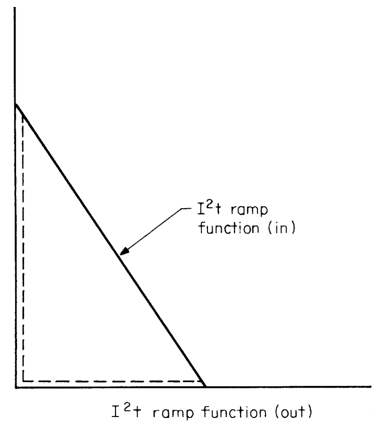


FIGURE 19-16 Sloping response curve for coordinating with thermal magnetic breakers or fuses.

conductors and load is a primary concern. Consideration is also given to the performance of circuit breakers with respect to other upstream and downstream circuit breakers and other protective devices. The object in coordinating protective devices is to make them selective in their operation with respect to each other. In doing so, the effects of short circuits on systems are reduced to a minimum by disconnecting only the affected parts of a system. Stated in another way, only the circuit breaker nearest the short circuit should open, leaving the rest of the system intact and able to supply power to the unaffected areas.

19.6 LOW-VOLTAGE BUSWAY

Low-Voltage Busways. Low-voltage busway (Fig. 19-17) provides a convenient, economical means of power distribution for a wide range of industrial and commercial buildings.

Description. Low-voltage (600 V and below) busways consist of a factory prefabricated housing which encloses current-carrying busbars. The housing protects the busbars from physical harm and provides mechanical support. Busbars, which are normally electroplated and insulated, may be either aluminum or copper.

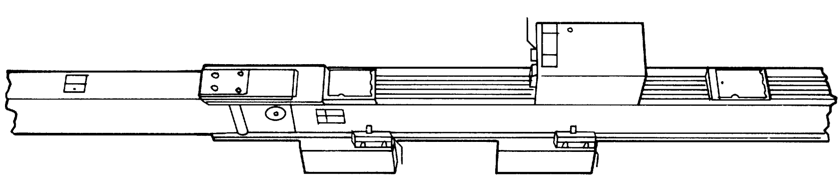


FIGURE 19-17 Low-voltage busway.

Types. Busway uses encompass a wide range of applications. For this reason, various types of busways have evolved. For medium- and high-current applications (100 to 5000 A), low-impedance *feeder* busway and *plug-in* busway are available. These types of busway are utilized to supply and/or distribute large blocks of power within a factory or office building. *Lighting* bus and *trolley* duct normally supply power to small lighting loads or moving loads such as hand tools on an assembly line. This type bus is rated 30 to 100 A.

Feeder busway is designed to provide maximum efficiency of power transfer from one point to another. If multiple-power tap-off points are required, then a plug-in busway is utilized. Plug-in busway is similar to a feeder busway except that plug-in openings are located at convenient intervals along the entire length. Various types of fusible or circuit breaker plug-in units may be easily attached at these openings to supply power to individual loads. Lighting contactors, motor starters, and other specialized devices are also available in plug-in unit form.

Applications. Busway uses are many and varied. However, they may be grouped into four major categories: (1) the main service busway, (2) the long feeder busway run, (3) the plug-in busway distribution, and (4) the electrical riser for tall buildings.

The *main service busway* is usually a fairly short run of high-ampere feeder busway which brings electricity from the power company supply point into the main switchboard in a building.

Long feeder busway runs are designed to carry large blocks of electrical power from one point to another with maximum efficiency. A typical example of this type of application would be a busway used to bring power from the main service switchboard to a remote distribution switchboard.

Plug-in busway distribution is the normal method for power distribution in most modern manufacturing areas today. The plug-in busway is positioned throughout the manufacturing area, and loads are connected to the busway using fusible or circuit breaker plug-in units. This wiring method is flexible and allows for the addition or movement of loads simply by relocating the plug-in units which attach to the busway.

Finally, *electrical busway risers* are used to form the backbone of the power distribution system in tall metropolitan buildings. The busway is positioned vertically within the building and tap-offs are used on each floor to provide the necessary power requirements. As load requirements increase in the future, the busway riser can be preplanned to allow for additional tap-offs, thus continuing to grow with the building.

Ratings. Low-voltage busway is available for applications up to 600 V ac and may be supplied for use on 3-phase, 3- or 4-wire systems. A wide range of feeder or plug-in busway ampere ratings may be selected (see Table 19-8). These ratings are based on actual heat rise within the busway. Current standards limit the maximum "hot-spot" temperature within the busway housing to 55°C above ambient.

TABLE 19-8 Low-Voltage Busway—Ampere Ratings

Plug-in		Feeder
■	225 A	
■	400 A	
■	600 A	
■	800 A	■
■	1000 A	■
■	1200 A	■
■	1350 A	■
■	1600 A	■
■	2000 A	■
■	2500 A	■
■	3000 A	■
■	4000 A	■
	5000 A	■

Ground Bus. A busway with a separate, low-impedance ground conductor is also available to provide for (1) a low-level static ground path, and (2) a secure low-impedance ground return path for medium- and high-level ground faults. This ground conductor is generally bonded to the busway housing at intervals of 10 ft or less. It is designed to carry a continuous 50% ground fault current as well as high-level ground fault currents of a short duration.

Conductors. A busway can be supplied with either aluminum or copper conductors. Based on size considerations only, copper is a more efficient electrical conductor than aluminum. However, when actual cost is considered, aluminum conductors are favored even though they must be slightly larger than the copper conductor required for the same ampere rating. Both aluminum and copper

TABLE 19-9 Low-Voltage Busway Short-Circuit Ratings

Continuous-Current rating of busway, A		Minimum short-circuit current ratings, A
Plug-in	Feeder	Symmetrical
100	. . .	10,000
225	. . .	14,000
400	. . .	22,000
600	. . .	22,000
	600	42,000
800	. . .	22,000
	800	42,000
1000	. . .	42,000
	1000	75,000
1200	. . .	42,000
1350	. . .	42,000
	1200	75,000
	1350	75,000
1600	. . .	65,000
	1600	100,000
2000	. . .	65,000
	2000	100,000
2500	. . .	65,000
	2500	150,000
3000	. . .	85,000
	3000	150,000
4000	. . .	85,000
	4000	200,000
	5000	200,000

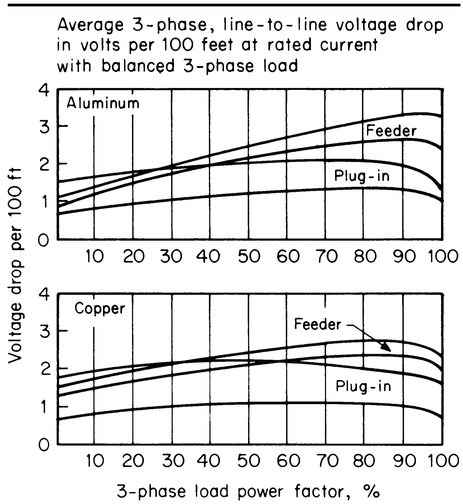
conductors are electroplated using tin or silver to ensure good surface conductivity at all electrical joints and plug-in unit connection points.

Available Fault Current Levels. In order to ensure that a busway can safely withstand the extreme physical forces, which can occur during a high-level fault, tests are conducted and maximum short-circuit levels are assigned to all busway components. The industry-recommended minimum short-circuit ratings for feeder and plug-in busway are shown in Table 19-9. In some instances optional higher short-circuit levels are available using alternate construction methods.

The actual available fault current throughout the entire electrical distribution system should be determined to ensure that all system components will function safely if a fault should occur.

Voltage Drop. Excessively low-voltage or wide-voltage variations on an electrical system can create many operating problems for the user. Modern low-impedance busway is designed to introduce an absolute minimum amount of voltage drop into the electrical system, thus helping to provide a stable source of electrical power. Typical feeder and plug-in busway voltage drop curves are shown in Table 19-10. Individual busway manufacturers can provide exact voltage drop curves for the specific busway ratings and types involved in each individual application.

TABLE 19-10 Typical Feeder and Plug-in Busway Voltage Drop Curves



Installation. A low-voltage busway system is made up of individual, factory-prefabricated components such as 10-ft straight lengths, elbows, tees, flanged ends, and cable tap boxes. These individual components are designed so they can be quickly joined together at the job site with a minimum of field labor. Typically, one grade-5 bolt and two large-diameter Belleville washers are utilized to connect all the bars in an entire bus stack (see Fig. 19-18). In some cases, special joint bolts are supplied with heads that pop off when torqued to the proper level, resulting in a tight connection. Larger ampere ratings may utilize two or three busbar stacks, which may require slight additional assembly time.

Prefabricated busway components reduce actual installation time to a minimum and result in a very low *installed cost* for the low-voltage busway. (The installed cost is the total material and labor cost necessary to provide a completed busway installation.)

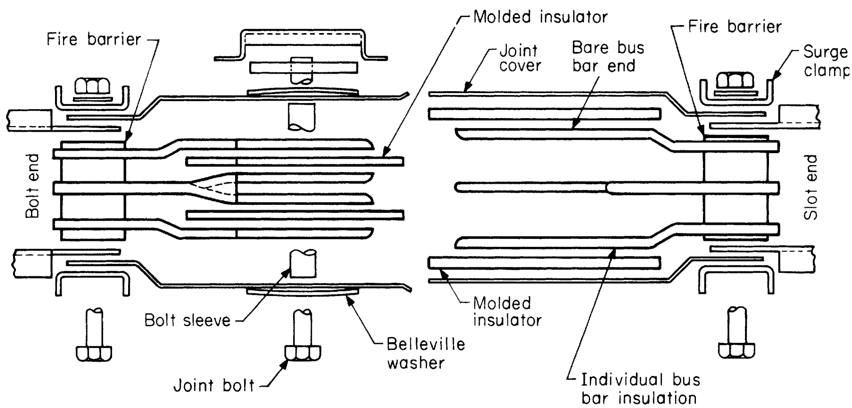


FIGURE 19-18 Busway one-bolt joint.

19.7 PROTECTIVE GROUNDING

Purpose of Grounding. Secondary ac distribution systems should be grounded at the neutral conductor if the maximum voltage to ground does not exceed 150 V and may be grounded if this voltage is above 150 but does not exceed 300 V (Fig. 19-19). This is to guard against imposition of a dangerous high voltage in case a breakdown in the transformer or crossing of primary- and secondary-circuit wires occurs.

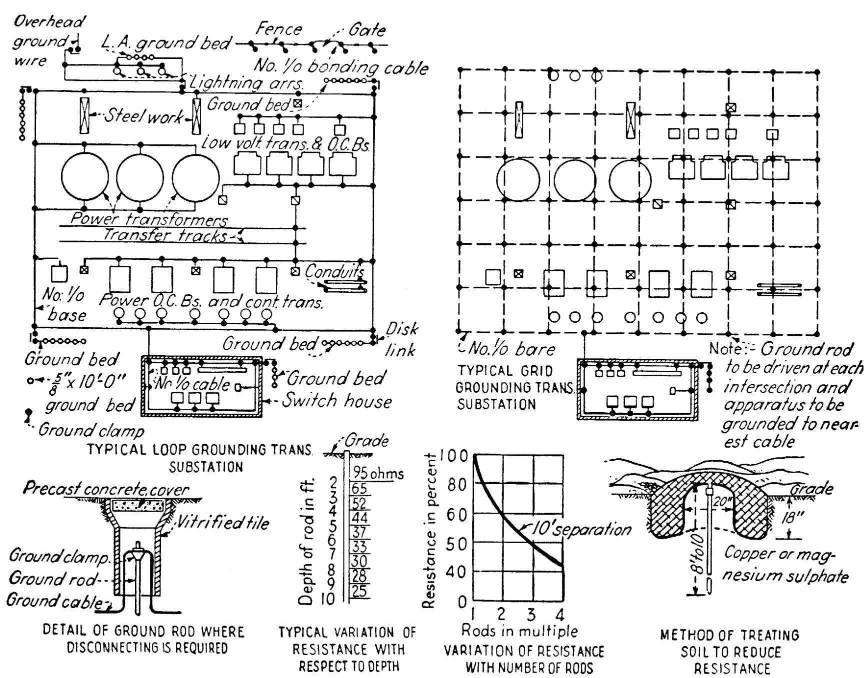


FIGURE 19-19 Complex equipment-grounding details showing connection, wire sizes, and grounds.

Conduits, metal raceways, cable armors, and metal cases or frames of equipment must be grounded (or isolated, as an alternative) and bonded so that, if this metal enclosure should come into contact with any of the circuit wires within it, no dangerous current would be passed to a person who touched the enclosure, since it is kept at ground potential.

The path to ground must be sufficiently low in resistance to allow enough current to flow to quickly open an overcurrent protective device and clear the fault. Otherwise, a voltage buildup will occur on metallic parts, with consequent hazards.

Size and Location of Ground Connection. The service neutral should be grounded at the point of entrance ahead of any disconnecting equipment with a copper wire or bus not smaller than indicated in Table 19-11. The neutral cannot be grounded at points beyond the load side of the disconnecting means (see NEC Article 250.24(A)(5)).

Metallic enclosures are commonly grounded through the grounding conductor used for the service neutral, although conduit or electrical metallic tubing may be used.

The grounding connection should be made to a grounding electrode such as (1) a metal underground water pipe, (2) a metal frame of a building, (3) concrete encased electrode, (4) a ground ring,

TABLE 19-11 Grounding Electrode Conductor for Alternating-Current Systems

Size of Largest Ungrounded Service-Entrance Conductor or Equivalent Area for Parallel Conductors* (AWG/kcmil)		Size of Grounding Electrode Conductor (AWG/kcmil)	
Copper	Aluminum or Copper-Clad Aluminum	Copper	Aluminum or Copper-Clad Aluminum†
2 or smaller	1/0 or smaller	8	6
1 or 1/0	2/0 or 3/0	6	4
2/0 or 3/0	4/0 or 250	4	2
Over 3/0 through 350	Over 250 through 500	2	1/0
Over 350 through 600	Over 500 through 900	1/0	3/0
Over 600 through 1100	Over 900 through 1750	2/0	4/0
Over 1100	Over 1750	3/0	250

Notes:

1. Where multiple sets of service-entrance conductors are used as permitted in NEC 230.40, exception no. 2, the equivalent size of the largest service-entrance conductor shall be determined by the largest sum of the areas of the corresponding conductors of each set.

2. Where there are no service-entrance conductors, the grounding electrode conductor size shall be determined by the equivalent size of the largest service-entrance conductor required for the load to be served.

*This table also applies to the derived conductors of separately derived ac systems.

†See installation restrictions in NEC 250.64(A).

(5) rod or pipe electrode or, (6) a plate electrode. See NEC Article 250.52. Where each type of electrode is present, they must be bonded together. See NEC Article 250.50. Resistance of a continuous underground water-piping system will usually be less than 0.1 Ω, thereby ensuring effectiveness.

Polarization of Wiring. The conductor to be grounded must be continuously identified throughout the system to avoid errors in connections. For No. 6 AWG or smaller, this is accomplished by finishing the insulating covering with a white or natural-gray finish. Ends of conductors larger than No. 6 AWG are painted white or gray or wrapped with white tape, where exposed in outlets or panelboards.

19.8 SYSTEMS OF INTERIOR DISTRIBUTION

Standard Secondary-Voltage Types. Common interior distribution systems for buildings or plants having appreciable loads are

1. 3-Phase 4-wire 120/208-V serving power and lighting.
2. Single-phase 3-wire 115/230-V serving lighting with 3-phase 3-wire 240- or 480-V for power.
3. 3-Phase 4-wire 277/480-V serving power and fluorescent or mercury lighting with single-phase 115/230-V circuits for other utilization provided from the power system by means of local dry-type transformers.

For very large buildings and industrial plants, distribution is often provided at higher voltage, notably 13.2, 4.1, or 2.3 kV, stepped down to utilization voltage at strategically load-centered substations.

3-Wire Single-Phase Systems. The 3-wire 115/230-V single-phase system (Fig. 19-20) is very commonly used for interior wiring for lighting and miscellaneous purposes but not for motor loads much in excess of 5 hp. The neutral wire is grounded; hence it should not be fused at any point. The branch circuits may be 2-wire 115- or 230-V or 3-wire 115/230-V. The neutral conductor carries only the unbalance in load between the two ungrounded conductors. For circuits up to 200 A, it should have the same capacity as the ungrounded conductors. A factor of 0.7 may be applied to unbalanced loads above 200 A in determining its size.

3-Phase 3-Wire Systems. These are usually employed where motors form a substantial load and where lighting is supplied from a separate single-phase system or by transformer. The usual voltage is 240 or 480 V. Branch circuits may be either 2-wire single-phase or 3-wire 3-phase.

3-Phase 4-Wire Systems. The 3-phase 4-wire system (Fig. 19-21) is widely used. Branch lighting circuits are connected between any one of the phase wires and the neutral wire. Power is taken from the 3-phase wires. The neutral wire is grounded. The voltage between phase wires is usually 208 V, and between any phase wire and neutral it is 120 V.

Circuits that may be supplied include 2-wire 120- or 208-V; 3-wire 120/208- or 208-V; and 4-wire 120/208-V. In the case of a 3-wire circuit consisting of 2-phase wires and the neutral, the "neutral" differs from that of a 3-wire single-phase circuit in that with both sides of the circuit evenly loaded it will carry equal current in the phase wires.

2-Wire Systems. The 2-wire system is used where current is supplied at 115 V, as when obtained from a 115-V generator, and on installations connected to utility companies' lines when the installation is of small capacity. It is limited in capacity; requires excessively large copper for heavy loads; and, in general, is not considered a modern complete distribution method.

2- and 3-Wire DC Systems. These systems are similar in connections (except grounding on the premises) and use to the 2- and 3-wire single-phase ac systems.

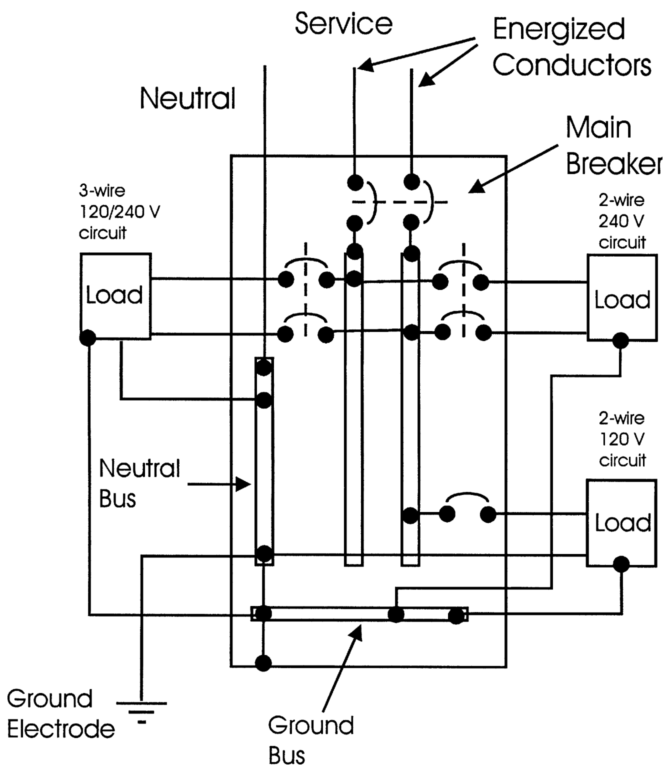


FIGURE 19-20 A 3-wire, 115/230-V, single-phase system.

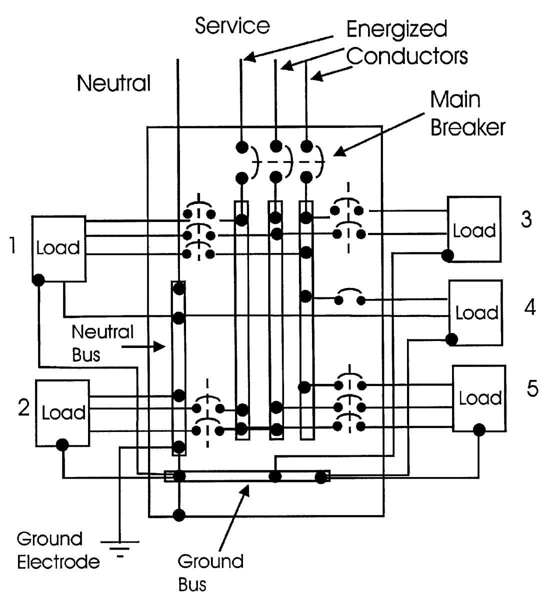


FIGURE 19-21 A 3-phase, 4-wire system: (1) 4-wire 120/208-V circuit; (2) 3-wire 120/208-V circuit; (3) 2-wire 208-V circuit; (4) 2-wire 120-V circuit; (5) 3-wire 208-V circuit.

2-Phase Distribution. This may be effected with 4 or 3 wires. In the former case, there is a pair of wires for each phase, while in the latter there is 1 wire for each phase and a common wire for both phases. The circuits must be balanced on either side, just as in the case of a 3-wire single-phase system. Where 3 wires are used, the common wire should be 1.4 times as large as either of the other two, since it must carry 1.4 times as much current. Motors are connected to both phases and employ all 3 or all 4 wires, as the case may be. With 4 wires, the lamps are connected to each phase as though the supply were single-phase, and care should be exercised to balance the phases as nearly as possible.

BIBLIOGRAPHY

Clapp, A. L., *National Electrical Safety Code Handbook*, Institute of Electrical and Electronics Engineers, New York, 1997.

Clark, W. H., *Electrical Design Guide for Commercial Buildings*, McGraw-Hill, New York, 1998.

McPartland, B. J., McPartland J. F., McPartland S. P., and Pullizzi J., *McGraw-Hill Handbook of Electrical Construction Calculations—Revised Edition*, McGraw-Hill, New York, 1998.

Miller, R., *Electrician's Pocket Manual*, McGraw-Hill, New York, 2005.

National Electrical Code®, NFPA 70, National Fire Protection Assoc., Quincy, MA, 2005.

Paschal, J., *EC&M's Electrical Calculations Handbook*, McGraw-Hill, New York, 2001.

Stallcup, J. G., *Stallcup's Electrical Design Book*, McGraw-Hill, New York, 2005.

Taylor, G., *Electrical Inspection Notes*, McGraw-Hill, New York, 2005.

SECTION 20
MOTORS AND DRIVES

Former contributors: Kenneth C. Cornelius, John H. Dulas, Alexander Kusko, Kelly A. Shaw, and Syed M. Peeran.

CONTENTS

20.1 GENERAL

Types of Electric Motors. Electric motors provide motive power to a wide variety of domestic and industrial machinery. Their versatility, reliability, and economy cannot be equaled by any other form of drive. Successful motor application depends on selecting a type of motor which satisfies the kinetic starting, running, and stopping requirements of the driven machinery. There are several methods of classifying electric motors. First, based on the electric power supply, motors are classified as dc and ac motors. Figure 20-1 shows further classification of ac and dc motors based upon the stator and rotor construction.

Classifications based upon size and applications are micro, fractional-horsepower, integral-horsepower, gear, torque, servo, and stepper motors in both standard and premium efficiency designs. Various types of enclosures have been standardized by the National Electric Manufacturers Association, U.S.A. (NEMA). The following are the standard enclosure types and their characteristics:

Types	Characteristics
Open:	
Dripproof	Operate with dripping liquids up to 15°C from vertical
Splashproof	Operate with splashing liquids up to 100°C from vertical
Guarded	Guarded by limited size openings (less than ¾ in)
Semiguarded	Only top half of motor guarded
Dripproof fully guarded	Dripproof motor with limited-size openings
Externally ventilated	Ventilated with separate motor-driven blower; can have other types of protection
Pipe ventilated	Openings accept inlet ducts or pipe for air cooling
Weather-protected type 1	Ventilating passages minimize entrance of rain, snow, and airborne particles; passages are less than ¾ in. in diameter
Weather-protected type 2	Motors have, in addition to type 1, passages to discharge high-velocity particles blown into the motor
Totally enclosed:	
Nonventilated (TENV)	Not equipped for external cooling
Fan-cooled (TEFC)	Cooled by external integral fan
Explosionproof	Withstands internal gas explosion; prevents ignition of external gas
Dust-ignitionproof	Excludes ignitable amounts of dust and amounts of dust that would degrade performance
Waterproof	Excludes leakage except around shaft
Pipe-ventilated	Openings accept inlet ducts or pipe for air cooling
Water-cooled	Cooled by circulating water
Water-and-air-cooled	Cooled by water-cooled air
Air-to-air-cooled	Cooled by air-cooled air
Guarded TEFC	Fan-cooled and guarded by limited-size openings
Encapsulated	Has resin-filled windings for severe operating conditions

NEMA classification according to the variability of speed includes constant-speed motors such as ac synchronous motors; induction motors with low, medium, or high slip; dc short-wound motors; varying-speed motors such as dc series motors or repulsion motors; and variable-speed motors such as dc shunt-, series-, and compound-wound motors.

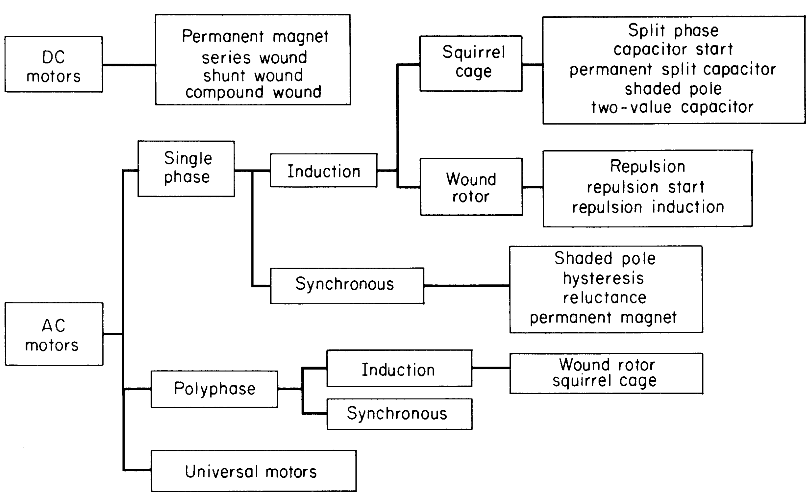


FIGURE 20-1 Classification of ac and dc motors.

Standards. Motors and generators are required to meet various industry and national standards and in some instances specific local codes and customer specifications. The more important of these standards may be briefly described as follows:

1. *NEMA Standards* are voluntary standards of the National Electrical Manufacturers Association and represent general practice in the industry. They define a product, process, or procedure with reference to nomenclature composition, construction, dimensions, tolerances, operating characteristics, performance, quality, rating, and testing. Specifically, they cover such matters as frame sizes, torque classifications, and basis of rating.
2. *IEEE Standards (AIEE)* concern fundamentals such as basic standards for temperature rise, rating methods, classification of insulating materials, and test codes.
3. *USA Standards* are national standards established by the United States of America Standards Institute, which represents manufacturers, distributors, consumers, and others concerned. USA Standards may be sponsored by any responsible body and may become national standards only if a consensus of those having substantial interest is reached. Standards may cover a wide variety of subjects such as dimensions, specifications of materials, methods of test, performance, and definition of terms. USA Standards frequently are those previously adopted by and sponsored by NEMA, IEEE, etc. The chief motor and generator standard of USASI is C50, "Rotating Machinery," which is substantially in agreement with current NEMA Standards.
4. *National Electrical Code* is a USA Standard sponsored by the National Fire Protection Association for the purpose of safeguarding persons and buildings from electrical hazards arising from the use of electricity for light, heat, power, and other purposes. It covers wiring methods and materials, protection of branch circuits, motors and control, grounding, and recommendations, regarding suitable equipment for each classification.
5. *Underwriters' Laboratories, Inc.* is an independent testing organization, which examines and tests devices, systems, and materials with particular reference to life, fire, and casualty hazards. It develops standards for motor and control for hazardous locations through cooperation with manufacturers. It has several different services by which a manufacturer can indicate compliance with Underwriters' Laboratories Standards. Such services are utilized on motors only in the case of explosionproof and dust-ignitionproof motors where label service is used to indicate to code-enforcing authorities that motors have been inspected to determine their adherence to Underwriters' Laboratories Standards for motors for hazardous locations.
6. *Federal Specification CC-M-641* for integral-horsepower ac motors has been issued by the federal government to cover standard motors for general government uses. Standard motors meet these specifications, but other Federal Specifications issued by various branches of the government for specific use may require special designs.
7. *World Standards*. Standards similar to our NEMA Standards have been established in other countries. The most significant are

   a. IEC (International Electrochemical Commission) Standard 72-1, Part 1

   b. German Standard DIN 42673

   c. British Standard BSI-2960, Part 2

   These standards specify dimensions, classes of insulation, and in some cases horsepower ratings.

20.2 DIRECT-CURRENT MOTORS

Classes of DC Motors. Direct-current motors are used in a wide variety of industrial applications because of the ease with which the speed can be controlled. The speed-torque characteristic may be varied to almost any useful form. Continuous operation over a speed range of 8:1 is possible. While ac motors tend to stall, dc motors can deliver over 5 times the rated torque (power supply permitting).

Reversal is possible without power switching. Permanent-magnet motors are available in fractional-horsepower ratings, while wound-field dc motors are classified as (1) shunt motor, in which the field winding is connected in parallel with the armature; (2) series motor, in which the field winding is connected in series with the armature; and (3) compound motor, which has a series-field and shunt-field winding. The shunt motor is used in constant-speed applications such as drives for dc generators in dc motor-generator sets. The series motor is used in applications where a high starting torque is required, such as in electric traction, cranes, and hoists. In compound motors, the droop of the speed-torque characteristic may be adjusted to suit the load.

The construction of dc motors with a wound field is practically identical to that of dc generators; with minor adjustment, the same dc machine may be operated either as a dc generator or as a motor. (See Sec. 8 of this handbook for construction, armature windings, commutator, etc.)

Permanent-magnet dc motors have fields supplied by permanent magnets that create two or more poles in the armature by passing magnetic flux through it. The magnetic flux causes the current-carrying armature conductors to create a torque. This flux remains basically constant at all motor speeds—the speed-torque and current-torque curves are linear.

Shunt Motors. DC shunt motors are suitable for application where constant speed is needed at any control setting or where appreciable speed range (by field control) is needed. The field circuit connection is shown in Fig. 20-2*a*.

Since a motor armature revolves in a magnetic field, an emf is generated in the conductors which is opposed to the direction of the current and is called the counter emf. The applied emf must be large enough to overcome the counter emf and also to send the armature current I_a through R_m, the resistance of the armature winding, the brushes; or

$$E_a = E_b + I_a R_m \qquad \text{volts} \tag{20-1}$$

where E_a = applied emf and E_b = counter emf. Since the counter emf at zero speed, that is, at starting, is identically zero and since normally the armature resistance is small, it is obvious in view of Eq. (20-1) that, unless measures are taken to reduce the applied voltage, excessive current will circulate in the motor during starting. Normally, starting devices consisting of variable series resistors are used to limit the starting current of motors.

The torque of a motor is proportional to the number of conductors on the armature, the current per conductor, and the total flux in the machine. The formula for torque is

$$\text{Torque} = 0.1175 Z\phi I_a \frac{\text{poles}}{\text{paths}} \times 10^{-8}\ \text{lb}\cdot\text{ft} \tag{20-2}$$

where Z = total number of armature conductors, ϕ = total flux per pole, and I_a = armature current taken from the line.

$$E_b = E_a - I_a R_m = Z\phi \frac{\text{r/min}}{60} \frac{\text{poles}}{\text{paths}} \times 10^{-8} \qquad \text{volts} \tag{20-3}$$

or

$$\text{r/min} = 60 \frac{E_a - I_a R_m}{Z\phi} \frac{\text{paths}}{\text{poles}} \times 10^{8} \tag{20-4}$$

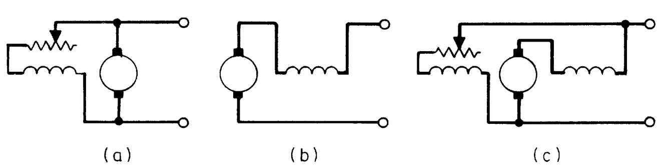


FIGURE 20-2 Field circuit connections of dc motor.

For a given motor, the number of armature conductors Z, the number of poles, and the number of armature paths are constant. The torque can therefore be expressed as

$$\text{Torque} = \text{constant} \times \phi I_a \qquad (20\text{-}5)$$

and the speed, likewise, is expressed as

$$\text{Speed} = \text{constant} \times (E_a - I_a R_m)/\phi \qquad (20\text{-}6)$$

In the case of the shunt motor, E_a, R_m, and ϕ are constant, and the speed and torque curves are shown as curves 1 (Fig. 20-3); the effective torque is less than that generated by the torque required for the windage and the bearing and brush friction. The drop in speed from no load to full load seldom exceeds 5%; indeed, since ϕ, the flux per pole, decreases with increase of load, owing to armature reaction, the speed may remain approximately constant up to full load.

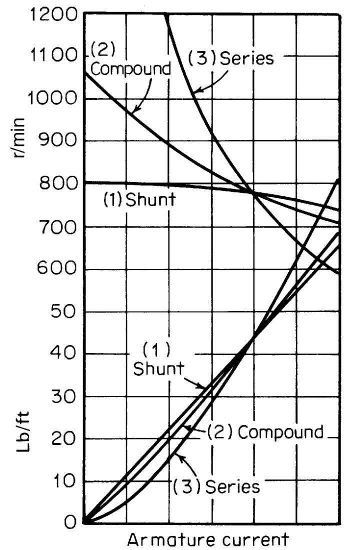


FIGURE 20-3 Motor characteristics.

Speed and Torque of Series Motors. Equations (20-6) and (20-5) apply to motors of all continuous-current types. In the case of series motors, the flux ϕ increases with the armature current I_a; the torque would be proportional to I_a^2 were it not that the magnetic circuit becomes saturated with increase of current. Since ϕ increases with load, the speed drops as the load increases. The speed and torque characteristics are shown in curves 3 (Fig. 20-3).

If the load on a series motor becomes small, the speed becomes very high, so that a series motor should always be geared or direct-connected to the load. If it were belted and the belt were to break, the motor would run away and would probably burst.

For a given load, and therefore for a given current, the speed of a series motor can be increased by shunting the series winding or by short-circuiting some of the series turns so as to reduce the flux. The speed can be decreased by inserting resistance in series with the armature.

Compound Motors. Compound-motor connections are shown in Fig. 20-2*c. The compound motor* is a compromise between the shunt and the series motors. Because of the series winding, which assists the shunt winding, the flux per pole increases with the load, so that the torque increases more rapidly and the speed decreases more rapidly than if the series winding were not connected; but the motor cannot run away under light loads, because of the shunt excitation. The speed and torque characteristics for such a machine are shown in curves 2 (Fig. 20-3).

The speed of a compound motor can be adjusted by armature and field rheostats, just as in the shunt machine.

Indirect compound is used on some dc motors. In this case, the heavy strap-wound series field is replaced by a wire-wound field similar to a small shunt field. This field is excited by an unsaturated dc exciter, usually separately driven at constant speed. This exciter is excited by the line current of the motor for which it supplies the series excitation (see Fig. 20-4). The output voltage and the current from the exciter are proportional to the main motor current; so a given proportionality exists between the load current of the motor and its wire-wound series-field strength. The use of a reversing switch and rheostat in the armature circuit of the series exciter permits variations in strength and even polarity

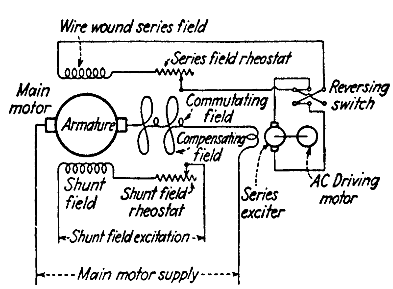


FIGURE 20-4 Direct-current motor with indirect compounding using a series exciter.

of the series field. This furnishes an easy method of changing the compounding of the motor if desired for various speeds, to maintain constant-speed regulation over a speed range. If desired, the series exciter rheostat can be mechanically connected to the shunt-field rheostat to accomplish this automatically.

Power Supplies. Power supplies to dc motors may be batteries, a dc generator, or rectifiers. The permanent-magnet and miniature motors use battery power supplies. Large integral-horsepower dc motors such as rolling-mill motors use dc generators as the power supply. Most fractional-horsepower and integral-horsepower dc motors operate with rectifier power supplies. Some of the types of rectifier power supplies are as follows:

1. Single-phase, half-wave
2. Single-phase, half-wave, back rectifier
3. Single-phase, half-wave, alternating-current voltage controlled
4. Single-phase, full-wave, firing angle controlled
5. Single-phase, full-wave, firing angle controlled, back rectifier
6. Three-phase, half-wave, voltage controlled
7. Three-phase, half-wave, firing angle controlled

The NEMA standard letter designations of dc motor test power supplies are as follows:

Power supply A—dc generator

Power supply C—3-phase 6-pulse controlled rectifier (230 V L-L, 60 Hz)

Power supply D—3-phase 6-pulse controlled rectifier (with three thyristors and three diodes) with free-wheeling diode (230/460 V L-L, 60 Hz)

Power supply E—3-phase 3-pulse controlled rectifier (460 V L-L, 60 Hz)

Power supply K—1-phase full-wave controlled rectifier with free-wheeling diode (230/115 V, 60 Hz)

When a direct-current integral-horsepower motor is operated from a rectified alternating-current supply, its performance may differ materially from that of the same motor when operated from a low-ripple direct-current source of supply, such as a generator or a battery. The pulsating voltage and current waveforms may increase temperature rise and noise and adversely affect commutation and efficiency. Because of these effects, direct-current motors must be designed or specially selected to operate on the particular type of rectified supply to be used. Armature-current form factor and ripple are two important parameters to be specified for motors which are required to operate with rectifier power supplies. The *form factor* is defined as the ratio of the rms value to the average value of the armature currents. Recommended rated form factors vary from 2.0 for 1-phase half-wave rectifier supplies to 1.1 for 3-phase full-wave rectifier supplies (see NEMA MG1-14.60). Because the letters used to identify the power supplies in common use have been chosen in alphabetical order of increasing magnitude of ripple current, a motor rated on the basis of one of these power supplies may be used on any power supply designed by a lower letter of the alphabet. For example, a motor rated on the basis of an E power supply may be used on a C or D power supply.

DC Motor Ratings. NEMA standard ratings of industrial dc motors for 240-V and 500/550-V dc supply voltages are given in Tables 10-4 and 10-5 of NEMA standard MG1. The rating is continuous unless otherwise specified. All short-term load tests shall commence only when the windings and other parts of the machine are within 5°C of the ambient temperature at the time of starting the test. Continuous and short-term ratings are based upon maximum ambient temperature and insulation class. Except in engine and boiler rooms, the maximum ambient temperature is 40°C and the insulation classes are A, B, and F, rated for temperature rises of 70°C, 100°C, and 130°C, respectively.

Losses and Efficiency. Power losses in dc motors are due to bearing friction, brush friction, windage, eddy currents and hysteresis in the armature core and pole faces, brush contact-drop, I^2R losses in the armature and field windings, and stray load losses. Typical values of total losses in industrial motors are 4% to 10% of the output. The bearing friction and brush friction losses are proportional to the speed of the motor, while the windage loss is proportional to the square of the speed. Eddy current loss in the armature teeth and in the armature core is proportional to the square of the speed and to the square of the air-gap flux density. Hysteresis loss in the armature teeth and core is proportional to the speed and the square of the flux density in the air gap. Brush contact drop is typically 1 V per brush arm for carbon-graphite brushes and 0.25 V for metal-graphite. Stray load losses are due to eddy currents in armature conductors, brush short-circuit losses in the commutator, and additional core loss arising from distortion of the magnetic field due to armature reaction. The efficiency of the dc motor is defined as

$$\eta = (\text{input electric power} - \text{losses})/\text{input power} \times 100\% \quad (20\text{-}7)$$

Typical efficiency variation with output is shown in Fig. 20-5.

Short-Time Ratings. The effect of time and enclosure on motor rating may be seen from the following: A given frame will have a rating of 12 hp at 500 r/min as an enclosed machine on continuous duty, or 19 hp at 500 r/min as an open machine on continuous duty, or 31 hp at 500 r/min with a 1-h rating, or 40 hp at 500 r/min with a $\frac{1}{2}$-h rating. The temperature rise on full load is 40°C as an open machine and 50°C as an enclosed machine. The horsepower is proportional to the speed over a range of 30% above or below the rated speed.

Methods of Speed Control. Speed of a dc motor is controlled either by varying the voltage across the armature, the field winding, or both. Series-parallel combinations are an effective means of reducing armature voltage and motor speed. This method is applied in cam-controlled traction motors. Two identical motors are connected in parallel or in series. When in parallel, full voltage is applied across each motor, causing it to run at base speed. When in series, the motor speeds are essentially one-half of base speed. Field-series resistance in shunt motors weakens the field, which causes the motors to run above the base speed. Speed range as high as 8:1 may be obtained in special motors. Armature-series resistance used with shunt or series motors produces motor speed below the base speed. In the series motor the field winding is also affected by the armature-series resistance, producing greater effect on the speed-torque characteristic than for the short motor where the field is constant. Speed control by this method is usually limited to approximately 50% of the base speed.

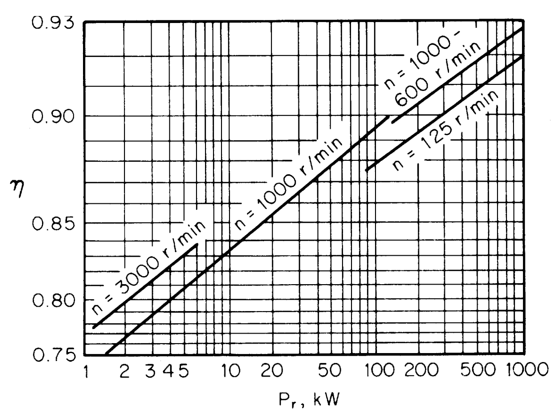


FIGURE 20-5 Typical efficiency curves of dc machines.

The above-speed control method results in power losses in the external resistors; solid-state dc motor control eliminates the power losses (see below).

Permanent-Magnet DC Motors. Permanent-magnet (PM) motors are available in fractional and low integral-horsepower sizes. They have several advantages over field-wound types. Excitation power supplies and associated wiring are not needed. Reliability is improved, since there are no exciting field coils to fail, and there is no likelihood of overspeed due to loss of field. Efficiency and cooling are improved by elimination of power loss in an exciting field. And the torque-versus-current characteristic is more nearly linear. Finally a PM motor may be used where a totally enclosed motor is required for a continuous-excitation duty cycle.

Temperature effects depend on the kind of magnet material used. Integral-horsepower motors with Alnico-type magnets are affected less by temperature than those with ceramic magnets because flux is constant. Ceramic magnets ordinarily used in fractional-horsepower motors have characteristics that vary about as much with temperature as do the shunt fields of excited machines.

Disadvantages are the absence of field control and special speed-torque characteristics. Overloads may cause partial demagnetization that changes motor speed and torque characteristics until magnetization is fully restored. Generally, an integral-horsepower PM motor is somewhat larger and more expensive than an equivalent shunt-wound motor, but total system cost may be less.

A PM motor is a compromise between compound-wound and series-wound motors. It has better starting torque, but approximately half the no-load speed of a series motor. In applications where compound motors are traditionally used, the PM motor could be considered where slightly higher efficiency and greater overload capacity are needed. In series-motor applications, cost consideration may influence the decision to switch. For example, in frame sizes under 5-in diameter the series motor is more economical. But in sizes larger than 5 in, the series motor costs more in high volumes. And the PM motor in these larger sizes challenges the series motor with its high torques and low no-load speed.

Brushless DC Motors. Brushless dc motors have a stationary armature and a rotating field structure, exactly opposite to how those elements are arranged in conventional dc motors. This construction speeds heat dissipation and reduces rotor inertia. Permanent magnets provide magnetic flux for the field. DC current to the armature is commutated with transistors rather than with the brushes and commutator bars of conventional dc motors.

Armatures of dc brushless motors typically contain 2 to 6 coils, whereas conventional dc motor armatures have from 10 to 50. Brushless motors have fewer coils because either two or four transistors are required to commutate each motor coil. This arrangement becomes increasingly costly and inefficient as the number of windings increases. A typical circuit of a brushless dc motor is shown in Fig. 20-6.

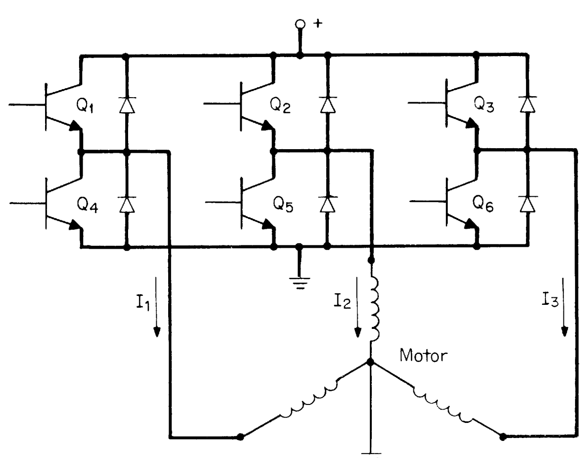


FIGURE 20-6 Typical circuit of a brushless dc motor.

The transistors controlling each winding of a dc brushless motor are turned on and off at specific rotor angles. The transistors provide current pulses to the armature windings that are similar to those provided by a commutator. The switching sequence is arranged to produce a rotating magnetic flux in the air gap that stays at a fixed angle to the flux produced by the permanent magnets on the rotor. Torque produced by a brushless dc motor is directly proportional to armature current.

DC Traction Motors. These are dc series motors typically rated 140 hp, 310 V, 2500 r/min. Four motors are used in each transit car, two on each axle. The power supply is 600 to 1000 V dc from the third rail, which is powered by 2500- to 5000-kW rectifier sets in rectifier substations located along the track. Starting and speed control are by either a cam controller or a chopper controller on board the transit car.

DC Servomotors. DC servomotors are high-performance motors normally used as prime movers in computers, numerically controlled machinery, or other applications where starts and stops must be made quickly and accurately. Servomotors have lightweight, low-inertia armatures that respond quickly to excitation-voltage changes. In addition, very low armature inductance in these motors results in a low electrical time constant (typically 0.05 to 1.5 ms) that further sharpens motor response to command signals. Servomotors include permanent-magnet, printed-circuit, and moving-coil (or shell) motors. The rotor of a shell motor consists of a cylindrical shell of copper or aluminum wire coils. The wire rotates in a magnetic field in the annular space between magnetic pole pieces and a stationary iron core. The field is provided by cast Alnico magnets whose magnetic axis is radial. The motor may have 2, 4, or 6 poles.

Each of these basic types has its own characteristics, such as inertia, physical shape, cost, shaft resonance, shaft configuration, speed, and weight. Although these motors have similar torque ratings, their physical and electrical constants vary considerably. The choice of a motor may be as simple as fitting one into the space available. However, this is generally not the case since most servosystems are very complex.

BIBLIOGRAPHY ON DC MOTORS

Anderson, E. P., *Electric Motors*, New York, Macmillan, 1991.

Beaty, H. W., and Kirtley, J. L., *Electric Motor Handbook*, New York, McGraw-Hill, 1998.

Chapman, S. J., *Electric Machinery Fundamentals*, New York, McGraw-Hill, 2005.

Dewan, S., Slemon, G. R., and Straughen, A., *Power Semi-Conductor Drives*, New York, Wiley, 1984.

Gotllieb, I. M., *Electric Motors and Control Techniques*, New York, McGraw-Hill, 1994.

Kusko, A., *Solid State—DC Motor Drives*, Cambridge, Mass., MIT Press, 1969.

Say, M. G., and Taylor, E. O., *Direct Current Machines*, New York, Wiley, 1980.

20.3 SYNCHRONOUS MOTORS

Definition. A synchronous motor is a machine that transforms electric power into mechanical power. That average speed of normal operation is exactly proportional to the frequency of the system to which it is connected. Unless otherwise stated, it is generally understood that a synchronous motor has field poles excited with direct current.

Types. The synchronous motor is built with one set of ac polyphase distributed windings, designated the *armature*, which is usually on the stator and is committed to the ac supply system. The configuration of the opposite member, usually the rotor, determines the type of synchronous motor. Motors with dc excited field windings on salient-pole or round rotors, rated 200 to 100,000 hp and larger, are the dominant industrial type. In the brushless synchronous motor, the excitation (field current) is supplied through shaft-mounted rectifiers from an ac exciter. In the slip-ring synchronous motor, the excitation is supplied from a shaft-mounted exciter or a separate dc power supply. Synchronous-induction motors rated below 5 hp, usually supplied from adjustable-speed drive inverters, are designed with a different reluctance across the air gap in the direct and quadrature axis to develop reluctance torque. The motors have no excitation source for synchronous operation. Synchronous motors below 1 hp usually employ a permanent-magnetic type of motor. These motors are usually driven by a transistor inverter from a dc source; they are termed *brushless dc motors*.

Standards. DC separately excited synchronous motors are covered by ANSI Standard C50.10-1965, Synchronous Machines, and C50.11-1965, Synchronous Motors. They are also covered by Part 21 of NEMA Standard MG-1 1972.

Theory of Operation. The operation of the dc separately excited synchronous motor can be explained in terms of the air-gap magnetic-field model, the circuit model, or the phasor diagram model of Fig. 20-7.

In the magnetic-field model of Fig. 20-7*a*, the stator windings are assumed to be connected to a polyphase source, so that the winding currents produce a rotating wave of current density J_a and radial armature reaction field B_a as explained below. The rotor carrying the main field poles is rotating in synchronism with these waves. The excited field poles produce a rotating wave of field B_d. The net magnetic field B_t is the spatial sum of B_a and B_d; it induces an air-gap voltage V_{ag} in the stator windings, nearly equal to the source voltage V_t. The current-density distribution J_a is shown for the current I_a in phase with the voltage V_t, and pf = 1. The electromagnetic torque acting between the rotor and the stator is produced by the interaction of the main field B_d and the stator

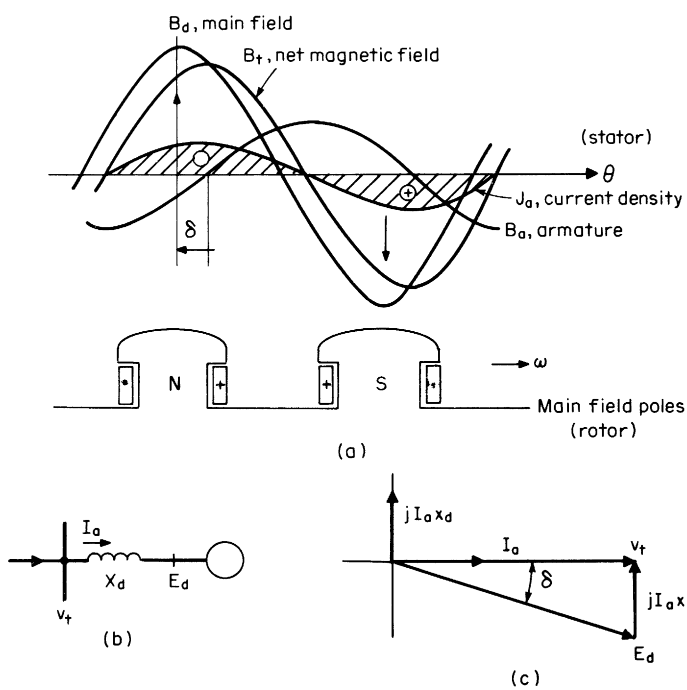


FIGURE 20-7 Operation of synchronous motor: (*a*) air-gap magnetic-field model; (*b*) circuit model; (*c*) phasor-diagram model.

current density J_a, as a $J \times B$ force on each unit volume of stator conductor. The force on the conductors is to the left $(-\phi)$; the reaction force on the rotor is to the right $(+\phi)$, and in the direction of rotation.

The operation of the synchronous motor can be represented by the circuit model of Fig. 20-7*b*. The motor is characterized by its synchronous reactance x_d and the excitation voltage E_d behind x_d. The model neglects saliency (poles), saturation, and armature resistance, and is suitable for first-order analysis, but not for calculation of specific operating points, losses, field current, and starting.

The phasor diagram of Fig. 20-7*c* is drawn for the field model and circuit model previously described. The phasor diagram neglects saliency and armature resistance. The phasors correspond to the waves in the field model. The terminal voltage V_t is generated by the field B_t; the excitation voltage E_d is generated by the main field B_d; the voltage drop $jI_a x_d$ is generated by the armature reaction field B_a; and the current I_a is the aggregate of the current-density wave J_a. The power angle δ is that between V_t and E_d, or between B_t and B_d. The excitation voltage E_d, in per-unit (pu),* is equal to the field current I_{fd}, in pu, on the air-gap line of the no-load (open-circuit) saturation curve of the machine.

Power-Factor Correction. Synchronous motors were first used because they were capable of raising the power factor of systems having large induction-motor loads. Now they are also used because they can maintain the terminal voltage on a weak system (high source impedance), they have lower cost, and they are more efficient than corresponding induction motors, particularly the low-speed motors. Synchronous motors are built for operation at pf = 1.0, or pf = 0.8 lead, the latter being higher in cost and slightly less efficient at full load.

The selection of a synchronous motor to correct an existing factor is merely a matter of bookkeeping of active and reactive power. The synchronous motor can be selected to correct the overall power factor to a given value, in which case it must also be large enough to accomplish its motoring functions; or it can be selected for its motoring function and required to provide the maximum correction that it can when operating at pf = 0.8 lead. In Fig. 20-8, a power diagram shows how the active and reactive power components P_s and Q_s of the synchronous motor are added to the components P_i and Q_i of an induction motor to obtain the total P_t and Q_t components, the kVA$_t$, and the power factor. The Q_s of the synchronous motor is based on the rated kVA and pf = 0.8 lead, rather than the actual operating kVA.

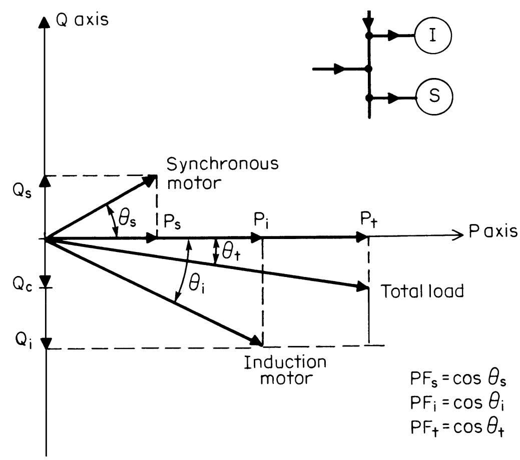


FIGURE 20-8 Power diagram of induction motor and synchronous motor operating in parallel, showing component and net values of *P* and *Q*.

The synchronous motor can support the voltage of a weak system, so that a larger-rating synchronous motor can be installed than an induction motor for the same source impedance. With an induction motor, both the *P* and *Q* components produce voltage drops in the source impedance. With a synchronous motor operating at leading power factor, the *P* component produces a voltage drop in the source resistance, but the *Q* component produces a voltage rise in the source reactance that can offset the drop and allow the terminal voltage to be normal. If necessary, the field current of the synchronous motor can be controlled by a voltage regulator connected to the motor bus. The leading current of a synchronous motor is able

*A reference unit for expressing all parameters on a common reference base. One pu is 100% of the chosen base.

to develop a sufficient voltage rise through the source reactance to overcome the voltage drop and maintain the motor voltage equal to the source voltage.

Starting. The interaction of the main field produced by the rotor and the armature current of the stator will produce a net average torque to drive the synchronous motor only when the rotor is revolving at speed n in synchronism with the line frequency f; $n = 120 f/p$, p = poles. The motor must be started by developing other than synchronous torques. Practically, the motor is equipped with an induction-motor-type squirrel-cage winding on the rotor, in the form of a damper winding, in order to start the motor.

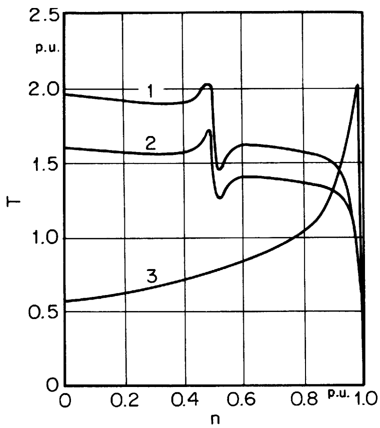


FIGURE 20-9 Characteristic torque curves for 5000-hp synchronous induction motor during runup at full voltage: (1) synchronous motor for pf = 1; (2) synchronous motor for pf = 0.8; (3) squirrel-cage induction motor.

The motor is started on the damper windings with the field winding short-circuited, or terminated in a resistor, to attenuate the high "transformer"-induced voltages. When the motor reaches the lowest slip speed, practically synchronous speed, the field current is applied to the field winding, and the rotor poles accelerate and pull into step with the synchronously rotating air-gap magnetic field. The damper windings see zero slip and carry no further current, unless the rotor oscillates with respect to the synchronous speed.

Starting curves for a synchronous motor are shown in Fig. 20-9. The damper winding is designed for high starting torque, as compared to an induction motor of the same rating. The closed field winding contributes to the starting torque in the manner of a 3-phase induction motor with a 1-phase rotor. The field winding produces positive torque to half speed, then negative torque to full speed, accounting for the anomaly at half speed. The maximum and minimum torque excursion at the anomaly is reduced by the resistance in the closed field winding circuit during starting. The effect is increased by the design of the damper winding.

The velocity of the rotor during the synchronizing phase, after field current is applied, is shown in Fig. 20-10. The rotor is assumed running at 0.05 pu slip on the damper winding. The undulation in speed, curve 1, is the effect of the poles attempting to synchronize the rotor just by reluctance torque. The added effect of the field current is shown by curve 2, and the resultant by curve 3. The effect of the reluctance torque of curve 1 is not dependent on pole polarity. The synchronizing torque of curve 2, with the field current applied, is pole polarity dependent; the poles want to match the air-gap

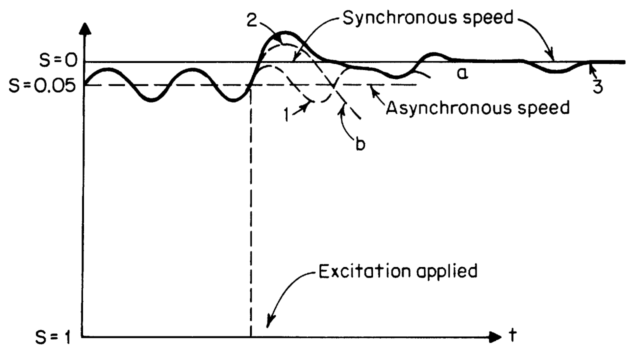


FIGURE 20-10 Relationship between slip and time for a synchronous motor pulling into synchronism: (*a*) successful; (*b*) unsuccessful.

TABLE 20-1 Locked-Rotor, Pull-In, and Pull-Out Torques for Synchronous Motors

r/min	hp	Locked rotor (Percent of rated full-load torque*)	Pull-in (based on normal Wk^2 of load)†	Pull-out† 1.0 pf	Pull-out† 0.8 pf
514 to 1800	200 and below; 1.0 pf 150 and below; 0.8 pf	100	100	150	175
	250 to 1000; 1.0 pf 200 to 1000; 0.8 pf	60	60	150	175
	1250 and larger	40	60	150	175
450 and below	All ratings	40	30	150	200

*The torque values with other than rated voltage applied are approximately equal to the rated voltage values multiplied by the ratio of the actual voltage to rated voltage in the case of the pull-out torque, and multiplied by the square of this ratio in the case of the locked-rotor and pull-in torque.

†With rated excitation current applied.

field in the forward torque direction. Curve *a* shows a successful synchronization. Curve *b* shows the condition of too much load or inertia to synchronize.

Torque Definitions. The torques described in the following paragraphs are listed in the Standards. The minimum values are given in Table 20-1.

Locked-rotor torque is the minimum torque, which the synchronous motor will develop at rest for all angular positions of the rotor, with rated voltage at rated frequency applied.

Pull-in torque is the maximum constant-load torque under which the motor will pull into synchronism, at rated voltage and frequency, when its rated field current is applied. Whether the motor can pull the load into step from the slip running on the damper windings depends on the speed-torque character of the load and the total inertia of the revolving parts. A typical relationship between maximum slip and percent of normal Wk^2 for pulling into step is shown in Fig. 20-11. Table 20-1 specifies minimum values of pull-in torque with the motor loaded with normal Wk^2; these values are given below. (See also Table 20-1.) *Nominal pull-in torque* is the value at 95% of synchronous speed, with rated voltage at rated frequency applied, when the motor is running on the damper windings.

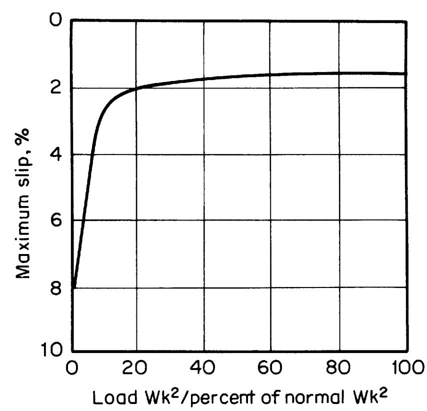


FIGURE 20-11 Typical relationship between load inertia and maximum slip for pulling synchronous motors into step.

Pull-out torque is the maximum sustained torque which the motor will develop at synchronous speed for 1 min, with rated voltage at rated frequency applied, and with rated field current.

In addition, the *pull-up torque* is defined as the minimum torque developed between standstill and the pull-in point. This torque must exceed the load torque by a sufficient margin to assure satisfactory acceleration of the load during starting.

The *reluctance torque* is a component of the total torque when the motor is operating synchronously. It results from the saliency of the poles and is a manifestation of the poles attempting to align themselves with the air-gap magnetic field. It can account for up to 30% of the pull-out torque.

The *synchronous torque* is the total steady-state torque available, with field excitation applied, to drive the motor and the load at synchronous speed. The maximum value as the motor is loaded is the pull-out torque, developed as a power angle $\delta = 90°$.

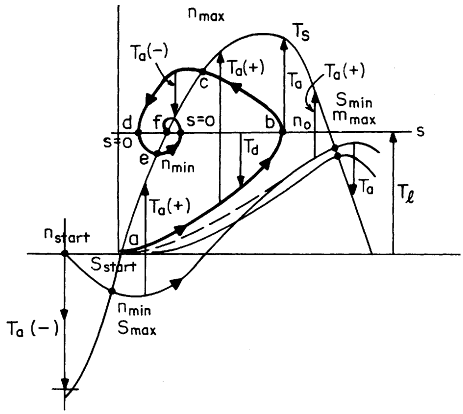


FIGURE 20-12 Locus of torque and speed versus power angle δ for a synchronous motor during a successful attempt and an unsuccessful attempt to synchronize.

Synchronization. Synchronization is the process by which the synchronous motor "pulls into step" during the starting process, when the field current is applied to the field winding. Initially, the rotor is revolving at a slip with respect to the synchronous speed of the air-gap magnetic-field waves. The rotor torque, produced by the damper windings, is in equilibrium with the load torque at that slip. The ability of the rotor to accelerate and synchronize depends upon the total inertia (Wk^2), the initial slip, and the closing angle of the poles with respect to the field wave at the instant field current is applied.

Figure 20-12 shows the torque versus angle δ locus for the rotor during a successful synchronization. The rotor is subjected to the synchronous torque T_s, which is a function of δ, and the damper torque T_d, which is a function of the slip velocity ($n_0 = n$). The torque T_a available to accelerate the rotor is the residual of $T_a = T_s + T_d - T_l$. In the figure, the closing angle is assumed zero at point *a*. Furthermore, $T_d = T_l$, so that the residual torque T_a is zero. The rotor has a finite slip, so that the power angle δ increases. As it does, the synchronous torque T_s increases, T_a increases and the rotor accelerates to point *b*, where $n = n_0$, $T_d = 0$. The slip goes negative, reverses the direction of the damper torque, but the rotor continues to accelerate to point *c*, where the speed is maximum and the accelerating torque is zero. The rotor falls back to points *d* and *e* at minimum speed, accelerates again, and finally synchronizes at point *f*.

If the initial slip is excessive, or if the inertia and/or load too great, the locus in Fig. 20-12 could follow the path *ab*′. The condition of $T_a = 0$ is reached below synchronous speed; the rotor never pulls into step, but oscillates around the initial slip velocity until the machine is tripped off.

Damper Windings. Damper windings are placed on the rotors of synchronous motors for two purposes: for starting and for reducing the amplitude of power-angle oscillation. The damper windings consist of copper or brass bars inserted through holes in the pole shoes and connected at the ends to rings to form the equivalent of a squirrel cage. The rings can extend between the poles to form a complete damper. Synchronous motors with solid pole shoes, or solid rotors, perform like motors with damper windings.

The design of the damper winding requires the selection of the bar and ring material to meet the torque and damping requirements. Figure 20-13 shows the effect on the starting curves for the damper winding of varying the material from a low-resistance copper in curve 1, to a higher-resistance brass or aluminum-bronze alloy in curve 2. Curve 1 gives a starting torque of about 0.25 pu, and a pull-in torque of 1.0 pu, of the nominal synchronous torque. Curve 2 gives a higher starting torque of about 0.5 to 1.0 pu, but a pull-in torque of about 0.4 pu of the nominal value. The additional starting torque of the field winding is superimposed on the torque of the damper alone. The damper winding must be designed to meet the characteristics of the load.

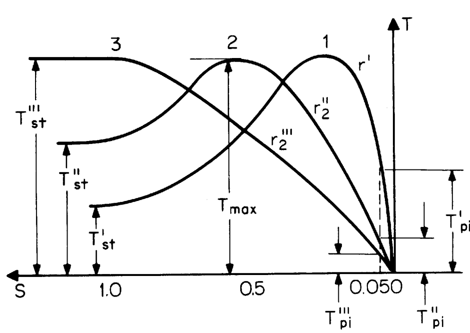


FIGURE 20-13 Effect of resistivity of damper material on the starting and pull-in torque of the synchronous motor. Damper winding 1, least resistance; damper winding 3, maximum resistance.

To design the damper winding so that the amplitude of the natural-frequency oscillation is reduced, the bar currents during the low-frequency sweeping of the air-gap flux across the pole faces must be maximized. Since the slip frequency is low, the currents and damper effectiveness are maximized by making the dampers low resistance, corresponding to curve 1 in Fig. 20-13. This design coincides with the requirement for low starting torque and high pull-in torque. In special cases, the equivalent of a deep-bar or double-bar damper can be used, if there is adequate space on the pole shoe.

Methods of Starting. The method used to start a synchronous motor depends on two factors: the required torque to start the load and the maximum starting current permitted from the line. Basically, the motor is started by using the damper windings to develop asynchronous torque or by using an auxiliary motor to bring the unloaded motor up to synchronous speed. Solid-state converters have also been used to bring up to speed large several-hundred-MVA synchronous motor/generators for pumped storage plants.

Techniques for asynchronous starting on the damper windings are the same as for squirrel-cage induction motors of equivalent rating. Across-the-line starting provides the maximum starting torque, but requires the maximum line current. The blocked-rotor kVA of synchronous motors as a function of pole number is shown in Fig. 20-14. If the ac line to the motor supplies other loads, the short-circuit kVA of the line must be at least 6 to 10 times the blocked rotor kVA of the motor to limit the line-voltage dip on starting. The starting and pull-in torques for three general classes of synchronous motors are shown in Fig. 20-15. The torques are shown for rated voltage; for across-the-line starting, the values will be reduced to V_t^2 (pu).

Reduced-voltage starting is used where the full starting torque of the motor is not required and/or the ac line cannot tolerate the full starting current. The starter includes a 3-phase open-delta or 3-winding autotransformer, which can be set to apply 50%, 65%, or 80% of line voltage to the motor on the first step. The corresponding torque is reduced to 25%, 42%, or 64%. The starter switches the motor to full voltage when it has reached nearly synchronous speed, and then applies the field excitation to synchronize the motor.

ANSI C50.11-1965 limits the number of starts for a synchronous motor, under its design conditions of Wk^2, load torque, nominal voltage, and starting method, to the following:

1. Two starts in succession, coasting to rest between starts, with the motor initially at ambient temperature, or
2. One start with the motor initially at a temperature not exceeding its rated load operating temperature.

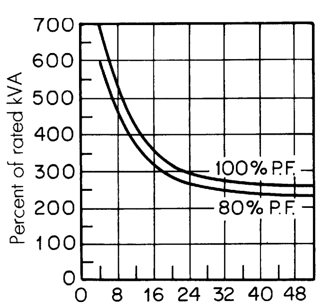


FIGURE 20-14 Approximate blocked-rotor kVA of synchronous motors.

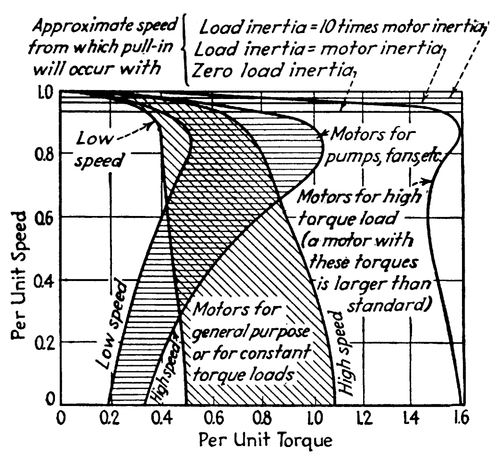


FIGURE 20-15 Approximate starting performance of synchronous motors.

If additional starts are required, it is recommended that none be made until all conditions affecting operation have been thoroughly investigated and the apparatus examined for evidence of excessive heating. It should be recognized that the number of starts should be kept to a minimum since the life of the motor is affected by the number of starts.

Exciters. DC separately excited synchronous motors are provided with a shaft-driven exciter to supply the field power. Exciters are classified into slip-ring types and brushless types. The slip-ring type consists of a dc generator, whose output is fed into the motor field winding through slip rings and stationary brushes. The brushless type consists of an ac generator, with rotating armature and stationary field; the output is rectified by solid-state rectifier elements mounted on the rotating structure and fed directly to the motor field winding. In each type, the motor field current is controlled by the exciter field current. Typical kilowatt ratings for exciters for 60-Hz synchronous motors are given in MG1-21.16 as a function of hp rating, speed, and power factor. For a given hp rating, the excitor kW increases as the speed is reduced, and as the power factor is shifted from pf = 1.0 to pf = 0.8 lead.

During starting, the motor field winding must be disconnected from the exciter and loaded with a resistor, to limit the high induced voltage, to prevent damage to the rectifier elements of the brushless type, and to prevent the circulation of ac current through a slip-ring-type dc exciter. The switching is done with a contactor for the slip-ring type, and with thyristors on the rotating rectifier assembly for the brushless type. Except for the disconnection for starting, the synchronous-motor excitation system is practically the same as for an ac generator of the same rating.

Brushless-type exciters are now used on all new high-speed synchronous motors (2 to 8 poles) that formerly were built with direct-drive dc exciters and slip rings. The brushless-type exciters require minimum maintenance and can be used in explosive-atmospheres. The circuit of a typical brushless-type excitation system is shown in Fig. 20-16. The semicontrolled bridge with three diodes and three thyristors rectifies the output of the ac exciter generator and supplies the motor field winding. The thyristors act as a switch to open the rectifier during starting and to close it during running, whereas the ac exciter generator is excited with its own field current. The resistor is permanently connected across the motor field winding during starting and running. It improves the torque characteristics during

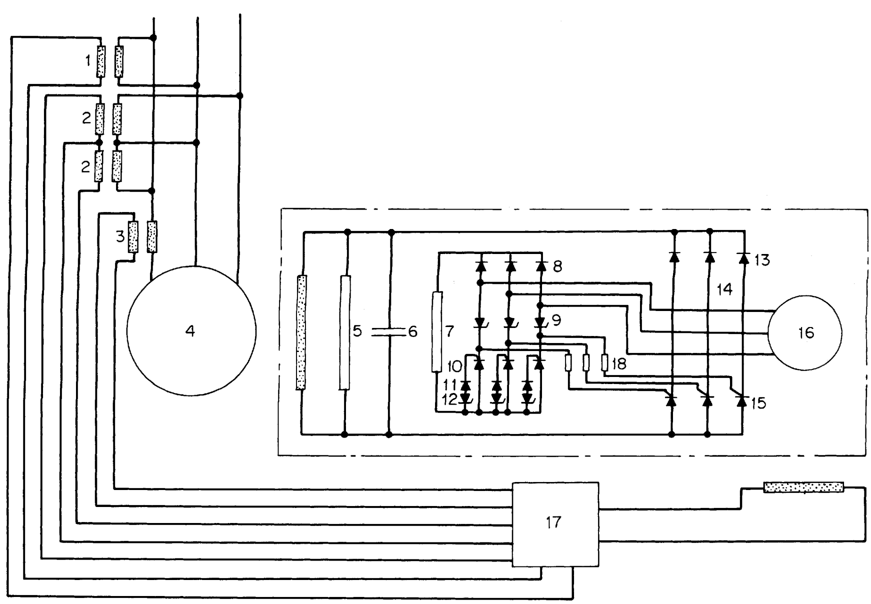


FIGURE 20-16 Brushless-type excitation system for a synchronous motor.

starting, and protects the bridge elements against transient overvoltages during running. The capacitor protects the diodes and thyristors against commutation overvoltages caused by hole-storage phenomena in conjunction with the inductances of the armature windings of the ac exciter generator.

The control system (Fig. 20-16) comprises a simple auxiliary rectifier arrangement connected in parallel with the main rectifier bridge and loaded with an auxiliary resistor 7. Each main thyristor has an auxiliary thyristor that provides the gate current and operates on the same phase of the ac excitor voltage. Consequently the trigger signal always occurs at the correct instant, that is, when the thyristors have a forward loading. No trigger signal is given during the blocking period. There is no excitation at the exciter during run-up, and therefore no trigger signal is applied to the gates of the thyristors and they remain blocking. The alternating current induced in the field winding flows in both directions through the protection resistor 5. When the machine has been run up to normal speed, the field voltage is applied to the ac exciter. It then supplies the control current and the thyristors are fired. Control losses are only 0.1% to 0.2% of the exciter power and are therefore negligible. The auxiliary thyristor 10 together with the diode 11 and Zener diode 12 prevents preignition of the thyristors during run-up due to high residual voltage in the ac exciter. On the other hand, the gates of the other thyristors are protected against overload by Zener diode 9 and resistor 18. If the voltage exceeds the Zener voltage, the Zener diode conducts the excess current.

Standard Ratings. Standard ratings for dc separately excited synchronous motors are given in NEMA MG1-1978, Part 21. Standard horsepowers range from 20 to 100,000 hp. Speed ratings extend from 3600 r/min (2-pole) to 80 r/min (90-pole) for 60-Hz machines, and five-sixths of the values for 50-Hz machines. The power factor shall be unity or 0.8 leading. The voltage ratings for 60-Hz motors are 200, 230, 460, 575, 2300, 4000, 4600, 6600, and 13,200 V. It is not practical to build motors of all horsepower ratings at these speeds and voltages.

Efficiency. Efficiency and losses shall be determined in accordance with IEEE test procedures for synchronous machines, Publication 115. The efficiency shall be determined at rated output, voltage, frequency, and power factor. The following losses shall be included in determining the efficiency: (1) I^2R loss of armature and field; (2) core loss; (3) stray-load loss; (4) friction and windage loss; and (5) exciter loss for shaft-driven exciter. The resistances should be corrected for temperature.

Typical synchronous motor efficiencies are shown in Fig. 20-17. The unity-power-factor synchronous motor—historically up to 3% more efficient than the NEMA design B induction motor, is now only 1% to 2% more efficient because of improvements in NEMA B designs and manufacturing techniques.

The 0.8 pf synchronous motor, because of the increased copper loss, is lower in efficiency; its efficiency is closer to that of the induction motor at high speed, but better at low speed.

Standard Tests. Tests on synchronous motors shall be made in accordance with *IEEE Test Procedure for Synchronous Machines*, Publ. No. 115, and ANSI C50.10-1965. The following tests shall be made on motors completely assembled in the factory and furnished with shaft and complete set of bearings: resistance test of armature and field windings; dielectric test of armature and field windings; mechanical balance; current balance at no load; and direction of rotation. The following tests may be specified on the same or duplicate motors: locked-rotor current; temperature rise; locked-rotor torque; overspeed; harmonic analysis and TIF; segregated losses; short-circuit tests at reduced voltage to determine reactances and time constants; field-winding impedance; and speed-torque curve.

The following tests shall be made on all motors not completely assembled in the factory: resistance and dielectric tests of armature and field windings. The following field tests are recommended after installation: resistance and dielectric tests of armature and field windings not completely assembled in the factory; mechanical balance; bearing insulation; current balance at no load; direction of rotation. The following field tests may be specified on the same or duplicate motors: temperature rise; short-circuit tests at reduced voltage to determine reactances and time constants; field-winding impedance.

The dielectric test for the armature winding shall be conducted for 1 min, with an ac rms voltage of 1000 V plus twice the rated voltage. For machines rated 6 kV and above, the test may be conducted with a dc voltage of 1.7 times the ac rms test value. The dielectric test for the field winding depends upon the connection for starting. For a short-circuited field winding, the ac rms test voltage is 10 times

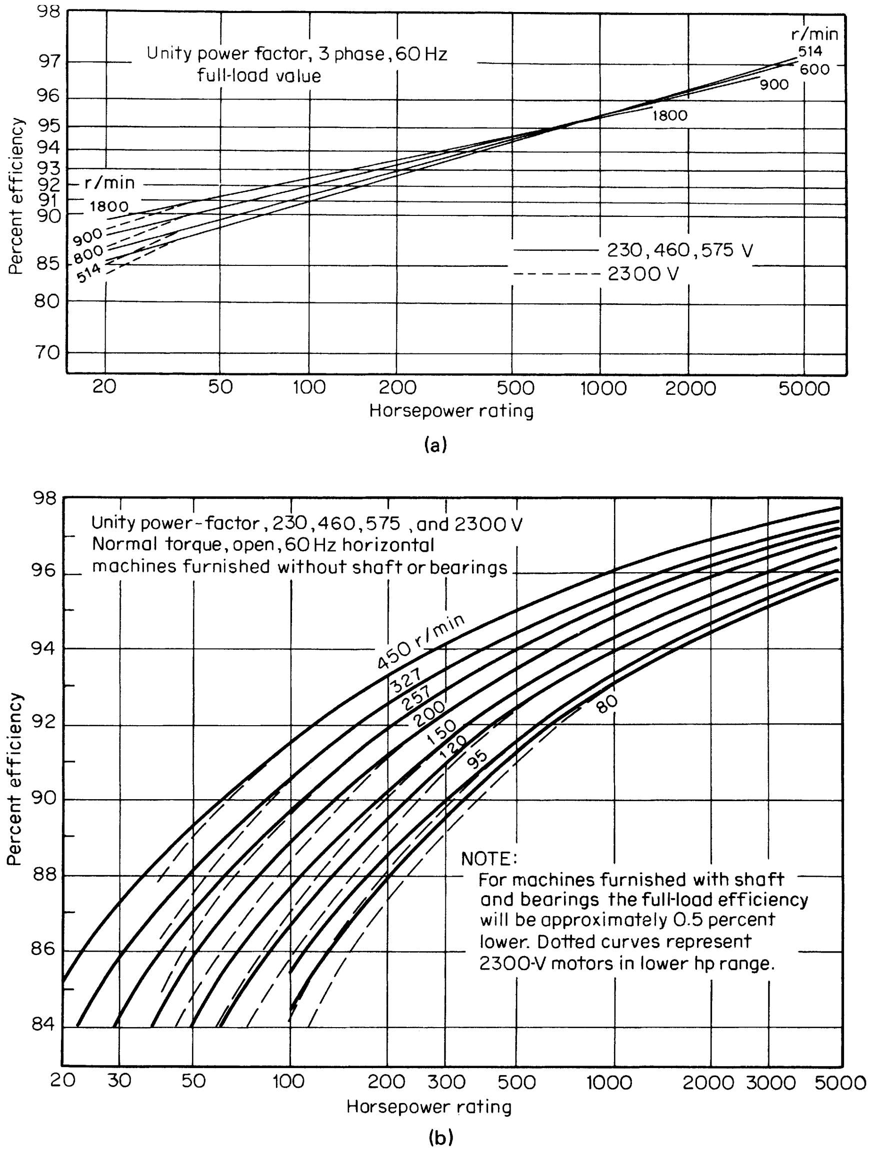


FIGURE 20-17 Full-load efficiencies of (*a*) high-speed general-purpose synchronous motors and (*b*) low-speed synchronous motors.

the rated excitation voltage, but no less than 2500 V, nor more than 5000 V. For a field winding closed through a resistor, the ac rms test voltage is twice the rms value of the *IR* drop, but not less than 2500 V, where the current is the value that would circulate with a short-circuited winding. When a test is made on an assembled group of several pieces of new apparatus, each of which has passed a high-potential test, the test voltage shall not exceed 85% of the lowest test voltage for any part of the group. When a test is made after installation of a new machine, which has passed its high-potential test at the factory and whose windings have not since been disturbed, the test voltages should be 75% of the original values.

Cycloconverter Drive. A unique application for large low-speed synchronous motors is for gearless ball-mill drives for the cement industry. For a recently installed drive, the motor is rated 8750 hp, 1.0 pf, 6850 kVA, 14.5 r/min 1900 V, 4.84 Hz, 40 poles, Class B. The power is provided by a cycloconverter over the range 0 to 4.84 Hz, as shown in Fig. 20-18. The cycloconverter consists of six thyristor rectifiers, each of which generates the polarity of the 3-phase ac voltage wave applied to the motor. The cycloconverter can be used effectively up to about one-third of the line frequency. The motor can be controlled in speed by the cycloconverter frequency, or in torque by the angle between the armature voltage and the field-pole position, approximately the power angle δ.

Inverter-Synchronous Motor Drive. Synchronous motors over about 1000 hp are being driven by machine-commutated inverters for adjustable-speed drives for large fans, pumps, and other loads. The machine-commutated inverter drive consists of two converters interconnected by a dc link as shown in Fig. 20-19*a*. The synchronous motor operates at constant volts per hertz, that is, voltage proportional to frequency and speed. The converter characteristics are shown in Fig. 20-19*b* and *c*. The $\pm V_d$ values are 1.35 times the line-line voltage on the ac side of each converter. For a given motor speed, frequency, and voltage, the firing angle of the rectifier is set at α_r to yield the required

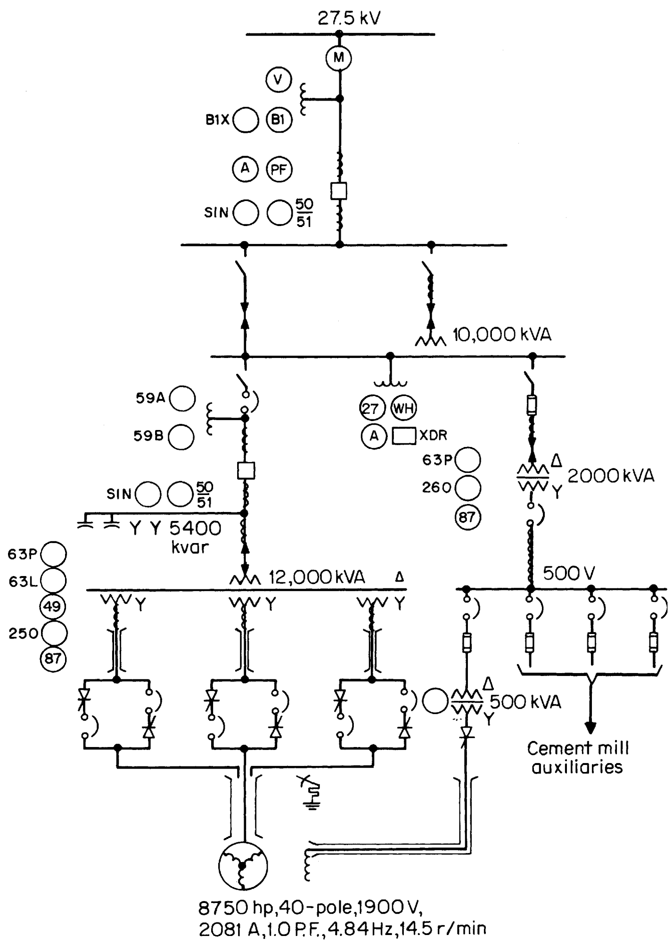


FIGURE 20-18 Cycloconverter-synchronous motor gearless drive system for ball mill.

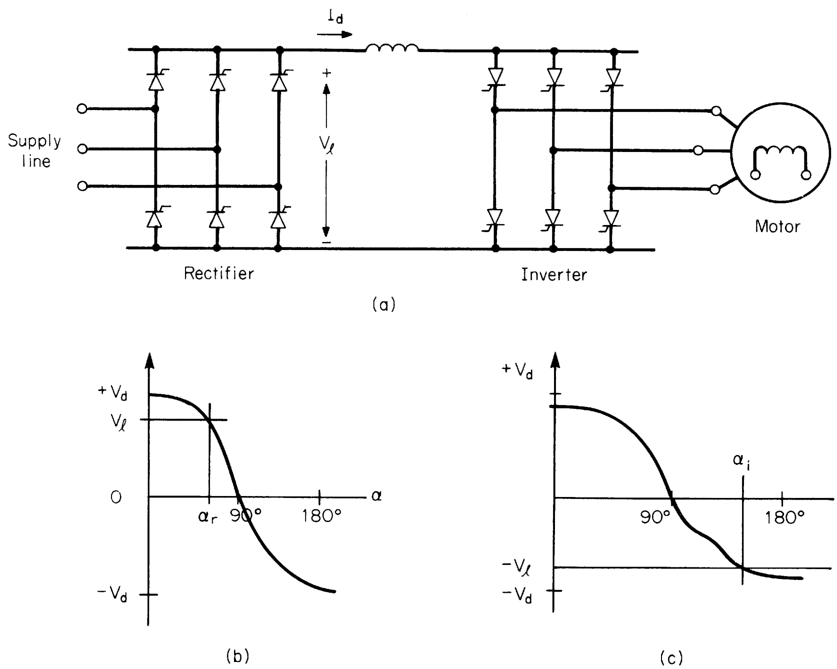


FIGURE 20-19 Diagram of (*a*) machine-commutated synchronous motor drive; (*b*) dc voltage vs. firing angle α_r, characteristic of rectifier; (*c*) dc voltage versus firing angle α_i characteristic of inverter.

dc voltage V_l for the link. The firing angle of the inverter is set at α_i in the inverting quadrant of the converter so that the link voltage V_l matches the internal ac voltage generated by the motor at the given speed. Power flows from the rectifier at $V_e I_d$ into the inverter and the motor. The inverter firing signals are synchronized to the motor voltage. For decelerating the motor, the rectifier and inverter functions are reversed by shifting the firing angles. Power flows from the motor into the dc link and to the supply line.

BIBLIOGRAPHY ON SYNCHRONOUS MOTORS

Beaty, H. W., and Kirtley, J. L., *Electric Motor Handbook*, New York, McGraw-Hill, 1998.
Fitzgerald, A. E., Kingsley, C., Jr., and Kusko, A., *Electric Machinery*, 3d ed., New York, McGraw-Hill, 1971.
IEEE Std. 115, *Test Procedures for Synchronous Machines.*
IEEE Std. 421, *Criteria and Definition for Excitation Systems for Synchronous Machines*.
Miller, T. J., *Brushless Permanent-Magnet and Reluctance Motor Drives*, Oxford University Press, 1989.
NEMA Std. MSI—*Motors and Generators*.

20.4 INDUCTION MACHINES

20.4.1 Theory of the Polyphase Induction Motor

Principle of Operation. An induction motor is simply an electric transformer whose magnetic circuit is separated by an air gap into two relatively movable portions, one carrying the primary and the other the secondary winding. Alternating current supplied to the primary winding from an electric power system induces an opposing current in the secondary winding, when the latter is short-circuited or closed through an external impedance. Relative motion between the primary and secondary structures is produced by the electromagnetic forces corresponding to the power thus transferred across the air gap by induction. The essential feature that distinguishes the induction machine from other types of electric motors is that the secondary currents are created solely by induction, as in a transformer, instead of being supplied by a dc exciter or other external power source, as in synchronous and dc machines.

Induction motors are classified as squirrel-cage motors and wound-rotor motors. The secondary windings on the rotors of squirrel-cage motors are assembled from conductor bars short-circuited by end rings or are cast in place from a conductive alloy. The secondary windings of wound-rotor motors are wound with discrete conductors with the same number of poles as the primary winding on the stator. The rotor windings are terminated on slip rings on the motor shaft. The windings can be short-circuited by brushes bearing on the slip rings, or they can be connected to resistors or solid-state converters for starting and speed control.

Construction Features. The normal structure of an induction motor consists of a cylindrical rotor carrying the secondary winding in slots on its outer periphery and an encircling annular core of laminated steel carrying the primary winding in slots on its inner periphery. The primary winding is commonly arranged for 3-phase power supply, with three sets of exactly similar multipolar coil groups spaced one-third of a pole pitch apart. The superposition of the three stationary, but alternating, magnetic fields produced by the 3-phase windings produces a sinusoidally distributed magnetic field revolving in synchronism with the power-supply frequency, the time of travel of the field crest from 1-phase winding to the next being fixed by the time interval between the reaching of their crest values by the corresponding phase currents. The direction of rotation is fixed by the time sequence of the currents in successive phase belts and so may be reversed by reversing the connections of one phase of a 2- or 3-phase motor.

Figure 20-20 shows the cross section of a typical polyphase induction motor, having in this case a 3-phase 4-pole primary winding with 36 stator and 28 rotor slots. The primary winding is composed

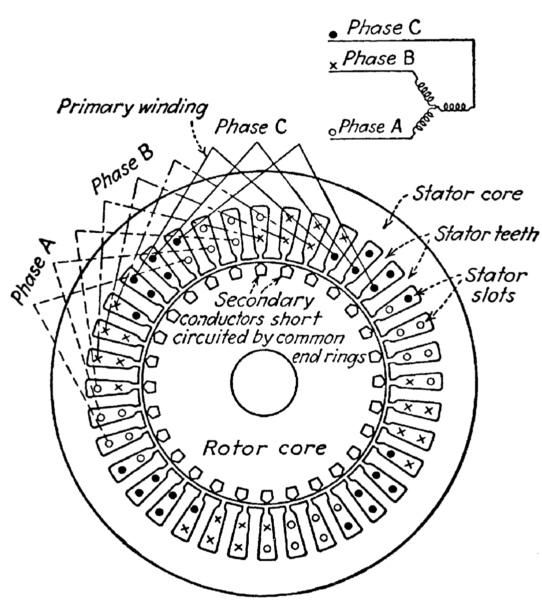
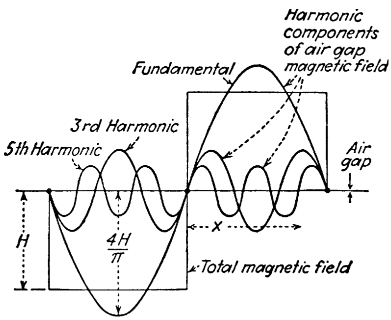


FIGURE 20-20 Section of squirrel-cage induction motor, 3-phase, 4-pole, 8/9-pitch stator winding.

of 36 identical coils, each spanning 8 teeth, one less than the 9 teeth in one pole pitch. The winding is therefore said to have $^8/_9$ pitch. As there are three primary slots per pole per phase, phase A comprises four equally spaced "phase belts," each consisting of three consecutive coils connected in series. Owing to the short pitch, the top and bottom coil sides of each phase overlap the next phase on either side. The rotor, or secondary, winding consists merely of 28 identical copper or cast-aluminum bars solidly connected to conducting end rings on each end, thus forming a "squirrel-cage" structure. Both rotor and stator cores are usually built on silicon-steel laminations, with partly closed slots, to obtain the greatest possible peripheral area for carrying magnetic flux across the air gap.



FIGUE 20-21 Magnetic field produced by a single coil.

The Revolving Field. The key to understanding the induction motor is a thorough comprehension of the revolving magnetic field.

The rectangular wave in Fig. 20-21 represents the mmf, or field distribution, produced by a single full-pitch coil, carrying H At. The air gap between stator and rotor is assumed to be uniform, and the effects of slot openings are neglected. To calculate the resultant field produced by the entire winding, it is most convenient to analyze the field of each single coil into its space-harmonic components, as indicated in Fig. 20-21 or expressed by the following equation:

$$H(x) = \frac{4H}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \frac{1}{7}\sin 7x + \cdots\right) \tag{20-8}$$

When two such fields produced by coils in adjacent slots are superposed, the two fundamental sine-wave components will be displaced by the slot angle θ, the third-harmonic components by the angle 3θ, the fifth harmonics by the angle 5θ, etc. Thus, the higher space-harmonic components in the resultant field are relatively much reduced as compared with the fundamental. By this effect of distributing the winding in several slots for each phase belt, and because of the further reductions due to fractional pitch and to phase connections, the space-harmonic fields in a normal motor are reduced to negligible values, leaving only the fundamental sine wave components to be considered in determining the operating characteristics.

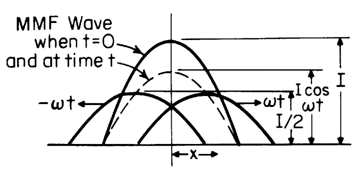


FIGURE 20-22 Resolution of alternating wave into two constant-magnitude waves revolving in opposite directions.

The alternating current flowing in the winding of each phase therefore produces a sine-wave distribution of magnetic flux around the periphery, stationary in space but varying sinusoidally in time in synchronism with the supply frequencies. Referring to Fig. 20-22, the field of phase A at an angular distance x from the phase axis may be represented as an alternating phasor $I \cos x \cos \omega t$ but may equally well be considered as the resultant of two phasors constant in magnitude but revolving in opposite directions at synchronous speed:

$$I \cos x \cos \omega t = \frac{I}{2}[\cos(x - \omega t) + \cos(x + \omega t)] \tag{20-9}$$

Each of the right-hand terms in this equation represents a sine-wave field revolving at the uniform rate of one pole pitch, or 180 elec deg, in the time of each half cycle of the supply frequency. The synchronous speed N_s of a motor is therefore given by

$$N_s = \frac{120f}{P} \quad \text{r/min} \tag{20-10}$$

where f = line frequency in hertz and P = number of poles of the winding.

Considering phase A alone (Fig. 20-23), two revolving fields will coincide along the phase center line at the instant its current is a maximum. One-third of a cycle later, each will have traveled 120 elec deg, one forward and the other backward, the former lining up with the axis of phase B and the latter with the axis of phase C. But at this moment, the current in phase B is a maximum, so that the forward-revolving B field coincides with the forward A field, and these two continue to revolve together. The backward B field is 240° behind the backward A field, and these two remain at this angle, as they continue to revolve. After another third of a cycle, the forward A and B fields will reach the phase C axis, at the same moment that phase C current becomes a maximum. Hence, the forward fields of all three phases are directly additive, and together they create a constant-magnitude sine-wave-shaped synchronously revolving field with a crest value two-thirds the maximum instantaneous value of the alternating field due to one phase alone. The backward-revolving fields of the three phases are separated by 120°, and their resultant is therefore zero so long as the 3-phase currents are balanced in both magnitude and phase.

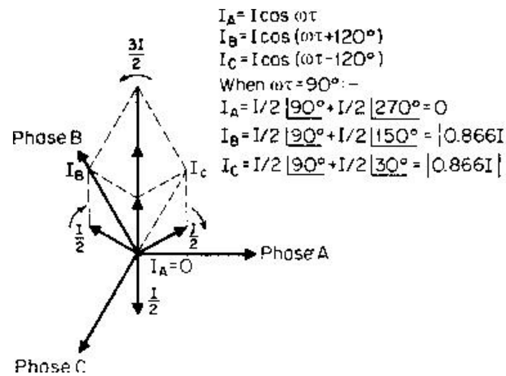


FIGURE 20-23 Resolution of alternating emf of each phase into oppositely revolving constant-magnitude components shown at instant when phase A current is zero ($\omega\tau = 90°$).

If a 2-phase motor is considered, it will have two 90° phase belts per pole instead of three 60° phase belts, and a similar analysis shows that it will have a forward-revolving constant-magnitude field with a crest value equal to the peak value of one phase alone and will have zero backward-revolving fundamental field. A single-phase motor will have equal forward and backward fields and so will have no tendency to start unless one of the fields is suppressed or modified in some way.

While the space-harmonic-field components are usually negligible in standard motors, it is important to the designer to recognize that there will always be residual harmonic-field values which may cause torque irregularities and extra losses if they are not minimized by an adequate number of slots and correct winding distribution. An analysis similar to that given for the fundamental field shows that in all cases the harmonic fields corresponding to the number of primary slots (seventh and nineteenth in a nine-slot-per-pole motor) are important and that the fifth and seventh harmonics on 3-phase, or third and fifth on 2-phase, may also be important.

The third-harmonic fields and all multiples of the third are zero in a 3-phase motor, since the mmf's of the three phases are 120° apart for both backward and forward components of all of them. Finally, therefore, a 3-phase motor has the following distinct fields:

1. The fundamental field with P poles revolving forward at speed N_s.
2. A fifth-harmonic field with $5P$ poles revolving backward at speed $N_s/5$.
3. A seventh-harmonic field with $7P$ poles revolving forward at speed $N_s/7$.
4. Similar thirteenth, nineteenth, twenty-fifth, etc., forward-revolving and eleventh, seventeenth, twenty-third, etc., backward-revolving harmonic fields.

Figure 20-24 shows a test speed-torque curve obtained on a 2-phase squirrel-cage induction motor with straight (unspiraled) slots. The torque dips due to three of the forward-revolving fields are clearly indicated.

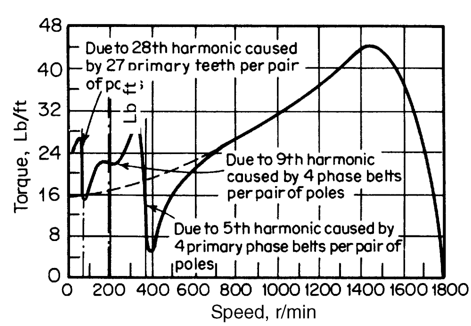


FIGURE 20-24 Speed-torque curve of 2-phase motor showing harmonic torque.

Torque, Slip, and Rotor Impedance. When the rotor is stationary, the revolving magnetic field cuts the short-circuited secondary conductors at synchronous speed and induces in them line-frequency currents. To supply the secondary *IR* voltage drop, there must be a component of voltage in time phase with the secondary current, and the secondary current, therefore, must lag in space position behind the revolving air-gap field. A torque is then produced corresponding to the product of the air-gap field by the secondary current times the sine of the angle of their space-phase displacement.

At standstill, the secondary current is equal to the air-gap voltage divided by the secondary impedance at line frequency, or

$$I_2 = \frac{E_2}{Z_2} = \frac{E_2}{R_2 + jX_2} \tag{20-11}$$

where R_2 = effective secondary resistance and X_2 = secondary leakage reactance at primary frequency.

The speed at which the magnetic field cuts the secondary conductors is equal to the difference between the synchronous speed and the actual rotor speed. The ratio of the speed of the field relative to the rotor to synchronous speed is called the slip *s*

$$s = \frac{N_s - N}{N_s}$$

or

$$N = (1 - s)N_s \tag{20-12}$$

where N = actual and N_s = synchronous rotor speed.

As the rotor speeds up, with a given air-gap field, the secondary induced voltage and frequency both decrease in proportion to *s*. Thus, the secondary voltage becomes sE_2, and the secondary impedance $R_2 + jsX_2$, or

$$I_2 = \frac{sE_2}{R_2 + jsX_2} = \frac{E_2}{(R_2/s) + jX_2} \tag{20-13}$$

The only way that the primary is affected by a change in the rotor speed, therefore, is that the secondary resistance as viewed from the primary varies inversely with the slip.

In practice, the effective secondary resistance and reactance, or R_2 and X_2, change with the secondary frequency, owing to the varying “skin effect,” or current shifting into the outer portion of the conductors, when the frequency is high. This effect is employed to make the resistance, and therefore the torque, higher at starting and low motor speeds, by providing a double cage, or deep-bar construction, as shown in Fig. 20-25. The leakage flux between the outer and inner bars makes the inner-bar reactance high, so that most of the current must flow in the outer bars or at the top of a deep bar at standstill, when frequency is high. At full speed, the secondary frequency is very low, and most of the current flows in the inner bars, or all over the cross section of a deep bar, owing to their lower resistance.

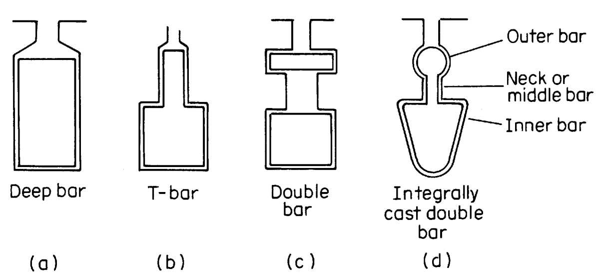


FIGURE 20-25 Alternative forms of squirrel cage rotor bars.

Analysis of Induction Motors. Induction motors are analyzed by two methods: (1) circle diagram and (2) equivalent circuit. The two methods are used for steady-state conditions. The circle diagram is convenient for visualizing overall performance but is too inaccurate for detailed calculations and design. The magnetizing current is not constant, but decreases with load because of the primary impedance drop. All of the circuit constants vary over the operating range due to magnetic saturation and skin effect. The equivalent circuit method predominates for analysis and design under steady-state conditions. The impedances can be adjusted to fit the conditions at each calculation point.

Circle Diagram. The voltage-current relations of the polyphase induction machine are roughly indicated by the circuit of Fig. 20-26. The magnetizing current I_M proportional to the voltage and lagging 90° in phase is nearly constant over the operating range, while the load current varies inversely with the sum of primary and secondary impedances. As the slip s increases, the load current and its angle of lag behind the voltage both increase, following a nearly circular locus. Thus, the circle diagram (Fig. 20-27) provides a picture of the motor behavior.

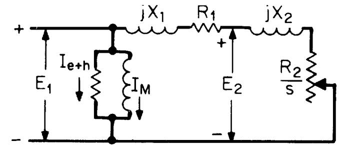


FIGURE 20-26 Equivalent circuit for circle diagram.

The data needed to construct the diagram are the magnitude of the no-load current *ON* and of the blocked-rotor current *OS* and their phase angles with reference to the line voltage *OE*. A circle with its center on the line *NU* at right angles to *OE* is drawn to pass through *N* and *S*. Each line on the diagram can be measured directly in amperes, but it also represents voltamperes or power, when multiplied by the phase voltage times number of phases. The line *VS* drawn parallel to *OE* represents the total motor power input with blocked rotor, and on the same scale *VT* represents the corresponding primary I^2R loss. Then *ST* represents the power input to the rotor at standstill, which, divided by the synchronous speed, gives the starting torque.

At any load point *A*, *OA* is the primary current, *NA* the secondary current, and *AF* the motor power input. The motor output power is *AB*, the torque X (synchronous speed) is *AC*, the secondary I^2R loss is *BC*, primary I^2R loss *CD*, and no-load copper loss plus core loss *DF*. The maximum power-factor point is *P*, located by drawing a tangent to the circle from *O*. The maximum output and maximum torque points are similarly located at *Q* and *R* by tangent lines parallel to *NS* and *NT*, respectively.

The diameter of the circle is equal to the voltage divided by the standstill reactance or to the blocked-rotor current value on the assumption of zero resistance in both windings. The maximum torque of the motor, measured in kilowatts at synchronous speed, is equal to a little less than the radius of the circle multiplied by the voltage *OE*.

Equivalent Circuit. Figure 20-28 shows the polyphase motor circuit usually employed for accurate work. The advantages of this circuit over the circle-diagram method are that it facilitates the derivation of simple formulas, charts, or computer programs for calculating torque, power factor, and

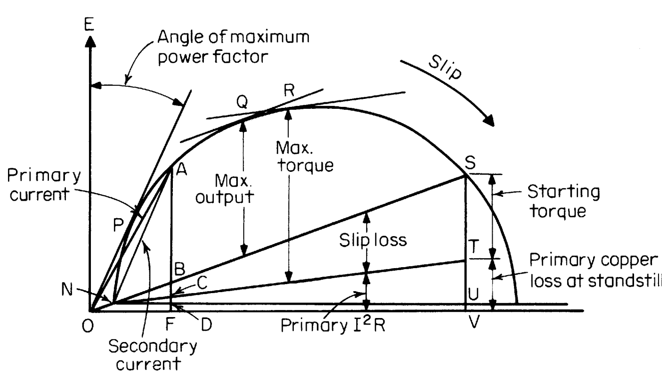


FIGURE 20-27 Circle diagram of polyphase induction motor.

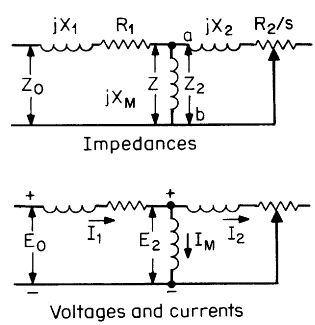


FIGURE 20-28 Equivalent circuit of polyphase induction motor.

TABLE 20-2 Formulas for Calculating Circuit Constants from Test Data for 3-Phase Motors

$$X = \frac{f}{f_t}\sqrt{\frac{V^2}{3I^2} - \left(\frac{W}{3I^2}\right)} \qquad \text{(see text)}$$

$X_1 = X_2 = 0.5X$ for single squirrel-cage or wound-rotor motors

$X_1 = 0.4X$ and $X_2 = 0.6X$ for low-starting-current motors

$W_H + W_F = W_{RL} - 3I_M^2 R_1$ (see text)

W_s (see text)

$$X_M = \frac{E_0}{I_M} - X_1$$

other motor characteristics and that it enables impedance changes due to saturation or multiple squirrel cages to be readily taken into account.

Formulas for calculating the constants from test data are given in Table 20-2, and their definitions are given in Table 20-3.

Inspection of the circuit reveals several simple relationships which are useful for estimating purposes. The maximum current occurs at standstill and is somewhat less than E/X. Maximum torque occurs when $s = R_2/X$, approximately, at which point the current is roughly 70% of the standstill current.

TABLE 20-3 Definitions of Equivalent-Circuit Constants

Unless otherwise noted, all quantities except watts, torque, and power output are per phase for 2-phase motors and per phase Y for 3-phase motors:

E_0 = impressed voltage (volts) = line voltage ÷ $\sqrt{3}$ for 3-phase motors
I_1 = primary current (amperes)
I_2 = secondary current in primary terms (amperes)
I_M = magnetizing current (amperes)
R_1 = primary resistance (ohms)
R_2 = secondary resistance in primary terms (ohms)
R_0 = resistance at primary terminals (ohms)
X_1 = primary leakage reactance (ohms)
X_2 = secondary leakage reactance (ohms)
$X = X_1 + X_2$
X_0 = reactance at primary terminals (ohms)
X_M = magnetizing reactance (ohms)
Z_1 = primary impedance (ohms)
Z_2 = secondary impedance in primary terms (ohms)
Z_0 = impedance at primary terminals (ohms)
Z = combined secondary and magnetizing impedance (ohms)
s = slip (expressed as a fraction of synchronous speed)
N = synchronous speed (revolutions per minute)
m = number of phases
f = rated frequency (hertz)
f_t = frequency used in locked-rotor test
T = torque (foot-pounds)
W_0 = watts input
W_H = core loss (watts)
W_F = friction and windage (watts)
W_{RL} = running light watts input
W_s = stray-load loss (watts)

Hence, the maximum torque is approximately equal to $E^2/2X$. This gives the basic rule that the percent maximum torque of a low-slip polyphase motor at a constant impressed voltage is about half the percent starting current.

By choosing the value of R_2, the slip at which maximum torque occurs can be fixed at any desired value. The maximum-torque value itself is affected, not by changes in R_2, but only by changes in X and to a slight degree by changes in X_M.

The magnetizing reactance X_M is usually 8 or more times as great as X, while R_1 and R_2 are usually much smaller than X, except in the case of special motors designed for frequent-starting service.

The equivalent circuit of Fig. 20-28 shows that the total power P_{g1} transferred across the air gap from the stator is

$$P_{g1} = mI_2^2 \frac{R_2}{s} \tag{20-14}$$

The total rotor copper loss is evidently

$$\text{Rotor copper loss} = mI_2^2 R_2 \tag{20-15}$$

The internal mechanical power P developed by the motor is therefore

$$\begin{aligned} P = P_{g1} - \text{rotor copper loss} &= mI_2^2 \frac{R_2}{s} - mI_2^2 R_2 \\ &= mI_2^2 R_2 \frac{1-s}{s} \qquad (20\text{-}16) \\ &= (1-s)P_{g1} \end{aligned}$$

We see, then, that of the total power delivered to the rotor, the fraction $1 - s$ is converted to mechanical power and the fraction s is dissipated as rotor-circuit copper loss. The internal mechanical power per stator phase is equal to the power absorbed by the resistance $R_2 (1 - s)/s$. The internal electromagnetic torque T corresponding to the internal power P can be obtained by recalling that mechanical power equals torque times angular velocity. Thus, when ω_s is the synchronous angular velocity of the rotor in mechanical radians per second

$$P = (1-s)\,\omega_s T \tag{20-17}$$

with T in newton-meters. By use of Eq. (20-16)

$$T = \frac{1}{\omega_s} mI_2^2 \frac{R_2}{s} \tag{20-18}$$

For T in foot-pounds and N_s in revolutions per minute

$$T = \frac{7.04}{N_s} mI_2^2 \frac{R_2}{s} \tag{20-19}$$

Torque and Power. Considerable simplification results from application of Thévenin's network theorem to the induction-motor equivalent circuit. Thévenin's theorem permits the replacement of any network of linear circuit elements and constant phasor voltage sources, as viewed from two terminals by a single phasor voltage source E in series with a single impedance Z. The voltage E is that appearing across terminals a and b of the original network when these terminals are open-circuited; the impedance Z is that viewed from the same terminals when all voltage sources within the network are short-circuited. For application to the induction-motor equivalent circuit, points a and b are taken as those so designated in Fig. 20-28. The equivalent circuit then assumes the forms given in Fig. 20-29. So far as phenomena to the right of points a and b are concerned, the circuits of Figs. 20-28 and 20-29 are identical when the voltage V_{1a} and the impedance $R_1 + jX_1$ have the proper values. According to

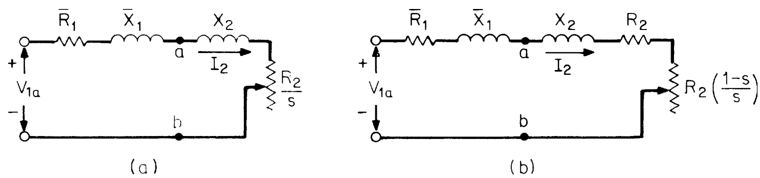


FIGURE 20-29 Induction-motor equivalent circuit simplified by Thévenin's theorem.

Thévenin's theorem, the equivalent source voltage V_{1a} is the voltage that would appear across terminals a and b of Fig. 20-28 with the rotor circuits open and is

$$V_{1a} = E_0 - I_0 (R_1 + jX_1) = E_0 \frac{jX_M}{R_1 + jX_{11}} \tag{20-20}$$

where I_M is the zero-load exciting current and

$$X_{11} = X_1 + X_M$$

is the self-reactance of the stator per phase and very nearly equals the reactive component of the zero-load motor impedance. For most induction motors, negligible error results from neglecting the stator resistance in Eq. (20-20). The Thévenin equivalent stator impedance $R_1 + jX_1$ is the impedance between terminals a and b of Fig. 20-28, viewed toward the source with the source voltage short-circuited, and therefore is

$$\overline{R}_1 + j\overline{X}_1 = R_1 + jX_1 \qquad \text{in parallel with} \quad jX_M$$

From the Thévenin equivalent circuit (Fig. 20-29) and the torque expression (Eq. 20-18), it can be seen that

$$T = \frac{1}{\omega_s} \frac{mV_{1a}^2(R_2/s)}{(\overline{R}1} + R_2/s)^2 + (\overline{X}_1 + X_2)^2 \tag{20-21}$$

The slip at maximum torque, $s_{\max T}$, is obtained by differentiating Eq. (20-21) with respect to s and equating to zero:

$$s_{\max T} = \frac{R_2}{\sqrt{\overline{R}_1^2 + (\overline{X}_1 + X_2)^2}}$$

The corresponding maximum torque is

$$T_{\max} = \frac{1}{\omega_s} \frac{0.5mV_{1a}^2}{\overline{R}_1 + \sqrt{\overline{R}_1^2 + (\overline{X}_1 + X_2)^2}}$$

20.4.2 Testing of Polyphase Induction Machines

Proof of guaranteed performance, the determination of torque or efficiency of driven machines, and the evaluation of design changes are some of the purposes that require accurate tests of induction machines. Normally, running-light, locked-rotor, resistance, and dielectric tests only are made on standard motors. Input-output tests or segregated-loss tests are made when accurate efficiency determination is required. The inconvenience of making input-output tests and the inaccuracies inherent in any method which determines the losses as a small difference between two large quantities make the segregated-loss methods of test preferable in many cases. Such tests are especially necessary when actual performance under the varying conditions of service is to be determined from a limited number of factory or laboratory test runs. Experience has shown that the equivalent-circuit method of calculation enables accurate predictions of efficiency and other performance data to be made, provided the circuit "constants" are determined in advance by careful tests.

The IEEE Test Code for Induction Machines[1] gives authoritative procedures for conducting all usual tests, and many of the data contained in the following sections are derived from this source.

Running-Light Test. The motor is run at no load with normal frequency and voltage applied, until the watts input becomes constant. On slip-ring motors, the brushes are short-circuited. Readings of amperes and watts are taken at one or more values of impressed voltage, with rated frequency maintained. Accurately balanced phase voltages and a sine-wave form of voltage are necessary for good results, requiring operation of the test alternator and transformers well below magnetic saturation. The watts input at rated voltage will be the sum of the friction and windage, core loss, and no-load primary I^2R loss. Subtracting the primary I^2R loss at the temperature of test from the input gives the sum of the friction and windage and core loss. Segregation of the core loss from the windage and friction is not necessary for normal efficiency or other rated-voltage performance calculations. However, the segregation can be made, if desired, by taking amperes and watts input readings, at rated frequency, at different voltages varying from 125% of normal down to about 15% voltage, or the point of minimum current. Plotting the input watts, less primary I^2R, against the square of the voltage and extrapolating the lower part of the curve in a straight line to intercept the zero-voltage axis determines the friction and windage. Typical data of such a test are shown in Fig. 20-30.

The value of the magnetizing reactance X_M in Fig. 20-28 is determined from the no-load current at rated voltage I_0 by using the value of primary leakage X_1 determined from locked-rotor test data.

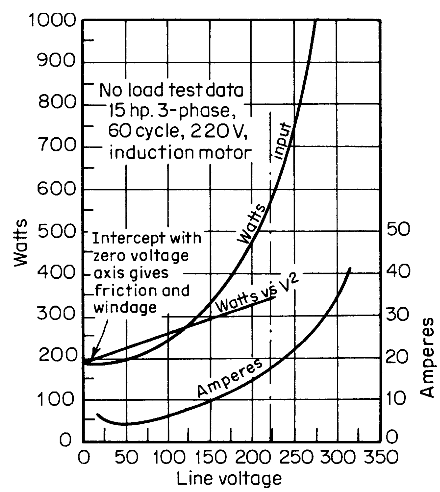


FIGURE 20-30 No-load excitation curves.

Locked-Rotor Test. The motor is blocked so it cannot rotate; a reduced voltage of rated frequency is applied to the terminals; and readings of volts, watts, and amperes are taken. Readings should be taken quickly, and the temperature of the windings should be observed before and after the test to minimize errors due to changing resistance values. In the case of machines with closed-slot rotors or very small air gaps, magnetic saturation of the leakage paths will occur, and it may be desirable to take readings at half or full voltage to establish the actual value of starting current. Equivalent-circuit performance calculations, however, should be based on data taken at approximately rated current.

When only low-voltage test data are available, the locked-rotor current at higher voltages can be estimated by the formula

$$I = \frac{V - V_0}{V_t - V_0} I_t \tag{20-22}$$

where V_t, I_t = test values of voltage and locked-rotor current; V, I = corresponding values at a different voltage; and V_0 = intercept of test current-voltage curve with zero-voltage axis, obtained by extrapolating the test curve as a straight line through points in the approximate range of 50% to 200% current. V_0 represents the voltage due to flux of saturation density crossing closed slot bridges and similar leakage flux paths.

The motor impedance per phase is determined from the volts, amperes, and watts readings. The total resistance component for a 3-phase motor is

$$R = \frac{W}{3I^2} \quad \Omega/\text{phase} \quad \text{Y} \tag{20-23}$$

and the reactance component is

$$X = \sqrt{\frac{V_2}{3I^2} - R^2} \quad \Omega/\text{phase} \quad \text{Y} \tag{20-24}$$

where W = watts input, I = line current, and V = voltage between lines.

Normally, the primary and secondary leakage-reactance values X_1 and X_2 are assumed equal, each having the value $X/2$.

The primary resistance is measured with direct current, a current about one-quarter of full-load value being preferably used, and readings being taken quickly to avoid errors due to temperature changes during the test. The primary resistance per phase Y is equal to one-half the resistance between any two terminals.

Subtracting the primary resistance at the temperature of test from the resistance component of the total impedance gives the effective secondary resistance at standstill. The starting torque may be calculated from this value by the equation

$$\text{Starting torque} = \frac{7.04KmI^2R_{2e}}{N_s} \quad \text{ft}\cdot\text{lb} \tag{20-25}$$

where I = amperes starting current per phase at specified voltage; m = number of phases; N_s = synchronous speed in r/min; R_{2e} = resistance component of motor impedance, less primary resistance at temperature of test, in ohms per phase; K = an empirical constant, usually approximately 0.9, which allows for nonfundamental secondary losses.

In practice, it is usual to measure the torque produced, by means of a lever arm and scale, in which case Eq. (20-25) provides a useful check on the accuracy of the measurements.

In the case of deep-bar or double squirrel-cage motors, the effective secondary reactance at speed is materially higher than at standstill, owing to the progressive shifting of the secondary current from the low-reactance, high-resistance paths into the low-resistance, high-reactance paths as the secondary, or slip, frequency decreases. Hence, for accurate performance calculations, it is necessary to determine the motor reactance at low secondary frequency. If a low-frequency supply is available, the locked-rotor test may be repeated at 15 Hz, or at most 25 Hz, for a 60-Hz motor. Calculation of the low-frequency reactance by Eq. (20-24) and multiplying this by the ratio of the rated to the test frequency will give the proper value to use in operating performance calculations.

Alternatively, the reactance value at speed may be obtained by adding an amount ΔX to the reactance determined by full-frequency locked-rotor test. The value of ΔX is approximately

$$\Delta X = R_{2e} - R_2 \tag{20-26}$$

where R_2 = secondary resistance of full-load slip, determined by the slip test.

Slip Test. Whenever feasible, a current-slip curve should be taken under actual load conditions, with rated voltage and frequency maintained at the motor terminals. Measurements at a few points in the neighborhood of full-load current are usually sufficient; but for slip-ring motors a wider range should be covered, owing to the variable resistance and should therefore be measured with a slip meter or stroboscopically. The slip-meter method makes use of a revolution counter differentially geared to the motor under test and to a small synchronous motor driven from the same power supply at the same synchronous speed. Care must be taken to correct the observed values of slip for the difference between the test temperature and the standard value of 75°C or the temperature attained in a full-load heat run with an ambient temperature of 25°C.

In practice, the value of current corresponding to an assumed value of $R_s 2/s$ is calculated exactly by the equivalent circuit; the corresponding value of s is read off the slip-current curve; and the true value of R_2 is obtained by multiplying R_2/s by this value of s. However, R_2 may be approximately determined as follows:

Very roughly, the secondary resistance is equal to

$$R_2 = 1.1\frac{E \cdot s}{I_1} \quad \text{approx} \cdot \Omega/\text{phase} \tag{20-27}$$

where E = terminal voltage per phase, s = ratio of revolutions per minute of slip to synchronous speed, and I_1 = observed phase current.

The coefficient 1.1 varies over a range of about 1 to 1.2, depending on the motor characteristics and the value of the test load.

In case direct slip measurements are not practicable, the value of R_2 determined by Eq. (20-23) in a low-frequency locked-rotor test may be used. Or, in the case of a wound rotor, the actual resistance between slip rings may be measured and multiplied by the square of the ratio of primary to secondary volts to obtain the resistance referred to primary. The voltage ratio is obtained by measurement of primary and secondary voltages at standstill with the slip rings open-circuited. Averages of several rotor positions are taken to avoid errors due to possible unbalance.

Stray-Load Loss Tests. Stray-load losses, W_s, are defined as the excess of the total measured losses above the sum of the friction and windage, core, and copper losses calculated for the conditions of load from the no-load tests described above. These extra losses are made up chiefly of high-frequency core losses and rotor I^2R losses caused by the pulsations of the leakage-reactance fluxes produced by load currents. While the stray-load losses may be determined by direct input-output tests with a dynamometer or calibrated driving motor, the result is a small difference between two large quantities and so accuracy is very difficult to obtain. Whenever such tests are made, it is desirable to repeat them with the direction of power flow reversed, so the measurement errors may be substantially canceled out.

There are several ways of determining the stray-load loss by separate loss measurements, but the procedure is fairly complex and must be carefully done if accurate results are to be obtained. These are described in the *IEEE Test Code for Polyphase Induction Machines*.

Performance Calculations. From the foregoing tests, all the circuit constants may be determined, enabling the equivalent-circuit calculations to be carried out. To facilitate this, the formulas for calculating the constants as defined in Table 20-3 are collected in Table 20-2.

The procedure in making performance calculations based on test data is first to divide E_0 by the appropriate expected value of normal current, an arbitrary value of R_2/s being thus obtained. With this value and the known circuit constants, calculations are carried through for one point, determining the actual value of I. By entering the test slip-current curve, the true value of s is found, and from this and R_2/s, R_2 is calculated. All the circuit constants are then known, whence the efficiency, power factor, torque, etc., are determined. Additional points are calculated with different values of s, covering the desired range of loads, and the exact characteristics are taken off curves plotted from the calculated results.

If values of torque, current, etc., are desired for considerable overloads or throughout the accelerating range, the value of R_2 and X should be modified to allow for magnetic saturation and eddy currents. Curves of reactance against current obtained by locked-rotor tests over the desired range of values and values of R_{2e} and corresponding values of ΔX obtained by locked-rotor tests at different frequencies are desirable for this purpose, especially in cases of closed-slot or double squirrel-cage rotors.

Temperature Tests. Temperature tests are made to determine the temperature rise of insulated windings under load conditions. ANSI Standards specify a limiting temperature for continuous-rated machines of 50°C by thermometer or 60°C by either the resistance- or the embedded-detector method for Class A insulating materials and corresponding values of 70°C by thermometer and 80°C by resistance or embedded detector for Class B insulation. Usually, the temperature is measured by mercury thermometers or thermocouples applied to the hottest accessible parts of the core and windings in several different locations. A small amount of putty is used to shield thermometer bulbs from the surrounding air, and care is taken to avoid external air currents, varying ambient temperatures, or other factors, which may introduce errors.

The preferred method of making a full-load temperature test is to maintain nameplate voltage, current, and frequency until the temperature becomes constant, readings being taken every half hour. When constant temperature is reached, the motor is stopped as quickly as possible and additional thermometers are applied to the rotating parts as soon as these have come to rest. The maximum permissible time of stopping is 1 min for machines of less than 50 kW rating, 2 min for 50 to 200 kW ratings, and 3 min for machines larger than 200 kW. The winding temperatures usually increase after shutdown; so readings must be recorded at frequent intervals until definitely falling temperatures are observed. The highest temperature reached at any time during the test is taken as the correct value. If the temperatures fall continuously after shutdown, a curve should be plotted of temperature versus time and extrapolated back to the moment of shutdown.

For protected-type or totally enclosed machines, it is often preferable to determine the temperature by the rise-of-resistance method. In this case, the "cold" resistance of the winding is measured at a known temperature, usually after the machine has been standing overnight at a uniform room temperature; and the "hot" resistance is measured immediately after shutdown. The hot resistance is taken as the highest value obtained after shutdown or is extrapolated back to the moment of shutdown if the resistance falls continuously.

The temperature is then calculated from the following formula:

$$T = \frac{R_T(234.5 + t)}{R_t} - 234.5 \qquad (20\text{-}28)$$

where T = winding temperature when R_T was measured, R_T = hot resistance, R_t = cold resistance, and t = winding temperature when R_t was measured.

Reference on Polyphase Induction Machine Testing

1. *American Standard Test Code for Polyphase Induction Motors and Generators*, USASC 50.20-1954.

20.4.3 Characteristics of Polyphase Induction Motors

Types. All polyphase induction motors may be classified as squirrel-cage or wound-rotor, and may be of the single-speed or multispeed type. Squirrel-cage motors are further classified by NEMA[1] for torque-speed and current-speed curves as Designs A, B, C, and D, and by Code designations from A to V for locked-rotor kVA/hp. For all induction motors, the allowable temperature rises and insulation systems are designated by classes A, B, F, and H. Finally, the mechanical dimensions are designated by frame sizes, and in enclosures from dripproof to totally enclosed with various types of ventilation. Both squirrel-cage and wound-rotor motors may be of the single-speed or multispeed type.

Based upon efficiency, motors are also classified as standard and energy-efficient motors. Several manufacturers have developed product lines of energy-efficient motors under various trade names. Some of these trade names are XE-Energy Efficient (Reliance Electric Co.), Energy Efficient Corro-Duty (US Electric Motors), and PE-21 Plus (Siemens).

Squirrel-Cage Motors. All integral-horsepower induction-motor design categories can mechanically withstand the magnetic stresses and locked-rotor torques of full-voltage line starting. The torque- and current-speed curves for Design A, B, C, and D squirrel-cage motors are shown in Figs. 20-31 to 20-33. Design B motors are most widely used; they have starting-torque and line-starting current characteristics suitable for most power systems. Design C and D motors have higher torque than Class B motors. For all design motors, the percentage torques tend to decline with

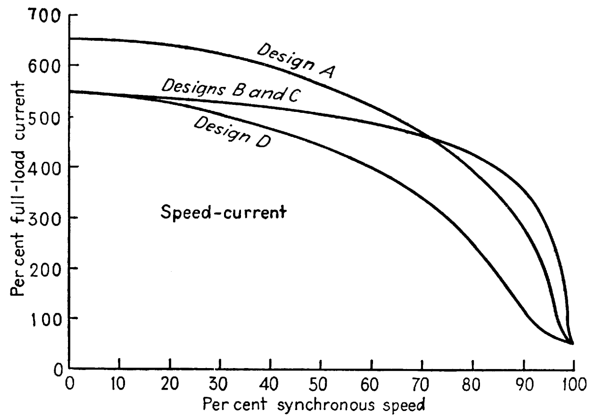


FIGURE 20-31 Typical speed-current curves for squirrel-cage induction motors.

increased hp rating cost. Typical load conditions and applications for Design A, B, C, and D motors are given in Table 20-4.

Wound-Rotor Motors. An insulated winding, usually 3 phase, is provided on the rotor; the terminal of each phase is connected to a slip ring on the shaft. The stationary brushes, which bear on the slip rings, are connected to external adjustable resistance or solid-state converters by which power can be removed from, or injected into, the rotor to adjust the speed. Speed-torque and speed-current curves for a typical wound-rotor motor for various amounts of external resistance are shown in Fig. 20-34. The numbers on the curves refer to the percent external resistance; 100% resistance gives rated torque at standstill. The use of solid-state converters in a modified Kraämer system is described later in reference to synchronous motor starting.

Wound-rotor motors are normally started with relatively high external resistance and this resistance is short-circuited in steps as the motor comes up to speed. Liquid rheostats are used in the higher ratings. This procedure allows the motor to deliver high-starting and accelerating torques, yet draw relatively light line current. Furthermore, most of the rotor-circuit losses during acceleration are dissipated in the external resistor rather than within the motor.

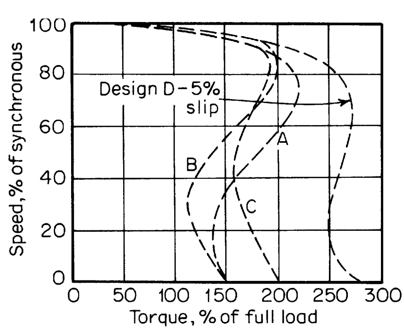


FIGURE 20-32 Speed-torque curves for typical NEMA standard Design A, B, C, and D squirrel-cage motors.

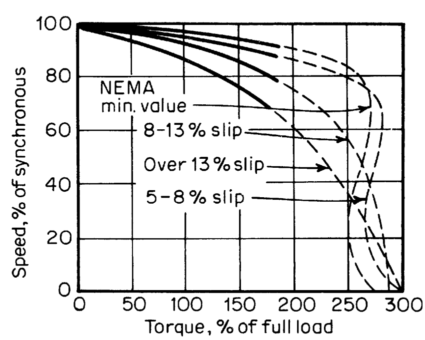


FIGURE 20-33 Speed-torque relationship for Design D squirrel-cage motors.

TABLE 20-4 3-Phase Electric Motor Selection Chart

For this type of equipment	Requiring these torques		With these load characteristics	Type and description
	Starting	Max. running		
Water supply pumps Industrial and chemical pumps Cooling towers Air-handling equipment Compressors Conveyors Process machinery Petroleum and chemical process equipment	100–150% of full-load torque	200-250% of full-load torque	Continuous operation, constant speed, high speed (over 720 r/min), easy starting; subject to short time overloads; good speed regulation	*Energy-efficient*: NEMA Design B, normal torques: normal starting current; can be used with variable-frequency variable-voltage inverters; higher efficiency than standard Design B motors
Centrifugal pumps Blower and fans Drilling machines Grinders Lathes Compressors Conveyors	100–150% of full-load torque	200–250% of full-load torque	Variable load conditions, constant speed; subject to short-time overloads; good speed regulation	*NEMA Design B:* normal torques: normal starting current; can be used with variable-frequency variable-voltage inverters
Reciprocating pumps Stokers Compressors Crushers Ball and rod mills	200–300% of full-load torque	Not more than full-load torque	High starting torque due to high inertia, back pressure, standstill friction, or similar mechanical conditions; torque requirements decrease during acceleration to full-load torque; not subject to severe overloads; good speed regulation	*NEMA Design C,* high torque: normal starting current; not recommended for use with variable-frequency inverters
Punch presses Cranes Hoists Press brakes Shears Oil-well pumps Centrifugals	Up to 300% of full-load torque	200–300% of full-load torque; loss of speed duringpeak loads required	Intermittent loads; may require frequent start, stop, and reverse cycles; machine uses a flywheel to carry peak loads; poor speed regulation to smooth power peaks; may require acceleration of high-inertia load	*NEMA Design D,* high torque: high slip; standard types have slip characteristics of 5–8% or 8–13% slip

Blowers Fans Machine tools Mixing machines Conveyors Pumps	Some require low torque; others require several times full-load torque	200% of full-load torque at each speed	Speed selection is desired, and two, three, or four fixed speeds are sufficient; starting torque can be low on blowers to high on conveyors; metal-cutting machines are usually constant hp; friction loads (conveyors) are usually constant torque; fluid or air loads (blowers) are variable torque	*Multispeed:* general normal torque on dominant winding or speed; consequent pole windings or separate windings for each speed; based on load requirement, can be constant horsepower, constant torque, variable torque
Crushers Conveyors Bending rolls Ball and rod mills Centrifugal blowers Pumps Printing presses Cranes and hoists Centrifugals	Can provide torque up to maximum torque at standstill	200–300% of full-load torque	Loads that require very high starting torque with low starting current; required speed adjustment over limited range (2 to 1); torque control during acceleration or controlled acceleration	*Wound-rotor:* requires rotor control system to provide desired characteristic; control may be resistors or reactors or fixed-frequency inverters in the secondary (rotor) circuit; actual load speed depends on setting of rotor control

Source: Andreas, J. C., *Energy Efficient Electric Motors,* 2d ed., New York, Marcel Dekker, 1992.

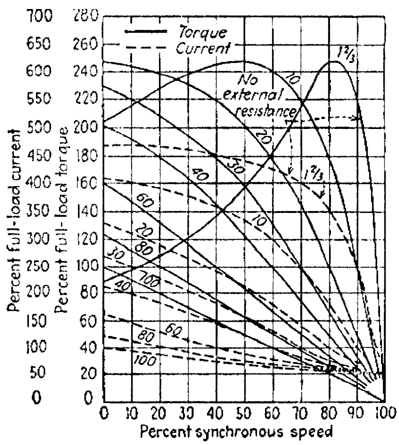


FIGURE 20-34 Speed-torque and speed-current curves of typical wound-rotor induction motor.

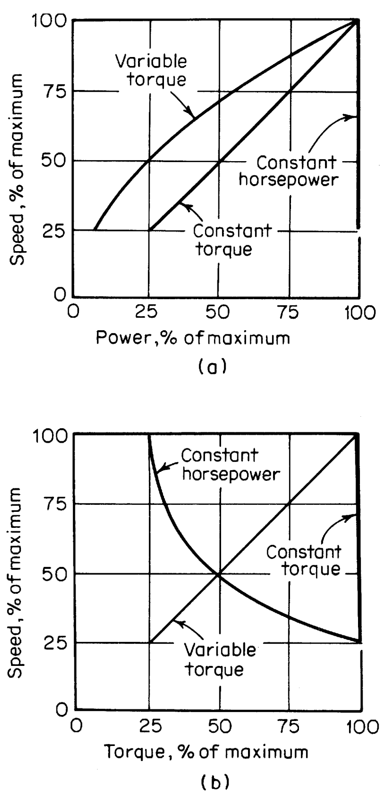


FIGURE 20-35 Basic load characteristics of multispeed motors having a 4:1 maximum speed ratio: (*a*) power; (*b*) torque.

The curves of Fig. 20-34 indicate that the external resistance reduces the speed at which the motor will operate with a given load torque. For any one value of external resistance, the motor has varying speed characteristics; any change in load results in a considerable change in speed. The lower the operating speed, the more pronounced the effect, so that it is usually not feasible to operate at less than 50% of full speed by this method. Furthermore, because the power loss in the rotor and external resistor is proportional to the slip, the efficiency is reduced in direct proportion to the speed reduction.

Breakdown torque is given by NEMA in MG1-12.40. Secondary data, including open-circuit slip-ring voltage and short-circuit slip-ring current, at standstill, are given in MG1-1034.

Slip-ring motors with external resistance are used as adjustable-speed motors from 50% to full speed for loads such as pumps and fans. They are used over the full speed range for hoists, elevators, and ski lifts. In addition, slip-ring motors are used to provide high starting and accelerating torque with low current for centrifuges, crushers, pulverizers, and other high-inertia loads. Solid-state ac and dc drives have replaced wound-rotor motors in many applications.

Multiple-Speed Squirrel-Cage Motors. Multispeed squirrel-cage motors may be of the single-winding or two-winding type. The former have a stator winding, which can be connected to give either one of two speeds having a ratio of 2:1. The method of connection is usually furnished by the controller manufacturer. The frame of the two-speed single-winding motor is about the same as that of the single-speed motor. The two-winding motor has two separate stator windings, which can be wound for any number of poles so that any two synchronous speeds can be obtained. In addition one or both of the stator windings may be arranged for reconnection as in a single-winding motor, giving a total of three or four speeds, but the two speeds obtained on a single winding must have a ratio of 2:1. Thus, a four-speed two-winding motor might have speeds of 1800, 900, 1200, and 600 r/min.

Multispeed motors are designed as (1) variable-torque motors, (2) constant-torque motors, and (3) constant-horsepower motors. The rated torque at four speed points for each type is shown in Fig. 20-35. Variable-torque motors have 1200/600 r/min, and are used on loads, such as in centrifugal pumps and fans whose horsepower requirement decreases more rapidly than the square of the reduction in speed. Constant-torque motors have horsepower ratings at each speed directly proportional to the speed, for example, 20/10 hp and 1200/600 r/min, and are used on conveyors, mixers, reciprocating compressors, printing presses, and other "constant-torque" loads. Constant-horsepower motors have the

TABLE 20-5 Temperature Rise for Single-Phase and Polyphase Induction Motors

	Class of insulation system			
	A	B	F	H
Integral horsepower				
All motors with 1.15 service factor or higher	70°C	90°C	115°C	—
Totally enclosed fan-cooled motors	60°C	80°C	105°C	125°C
Totally enclosed nonventilated motors	65°C	85°C	110°C	135°C
Motors with encapsulated windings, 1.0 service factor	65°C	85°C	110°C	—
All other motors	60°C	80°C	105°C	125°C
Fractional horsepower				
Open motors with 1.15 service factor or higher	70°C	90°C	115°C	—
Totally enclosed nonventilated and fan-cooled motors	65°C	85°C	110°C	135°C
Any motor in frame smaller than 42 frame	65°C	85°C	110°C	135°C
All other open motors	60°C	80°C	105°C	125°C

Note: Based on ambient temperature of 40°C, 3300-ft altitude. Temperature determined by the resistance method.

same horsepower rating at all speeds. They are used principally on machine tools, such as lathes, boring mills, planers, and radial drills. Multispeed motors of the constant-torque or variable-torque type are usually given a standard horsepower rating at the top speed but may have odd horsepower ratings at the lower speeds, since the latter are fixed by the speed ratios.

Temperature Rise. Temperature rise is no longer used as a rating method. Instead the manufacturer specifies the ambient temperature and the insulation class. The temperature rise will not exceed the limit for the insulation system when the motor is loaded to its rating or to its service factor load. The temperature rise limits are given in Table 20-5.

The temperature attained by squirrel-cage windings, cores, and mechanical parts shall not injure the machine in any respect. Temperatures shall be determined in accordance with the *IEEE Test Procedures*, Publication Nos. 112A and 114. For Class F and H insulation systems, special consideration shall be given to the bearings and lubrication.

The temperature rise for motors operating at any other ambient temperature T_a than 40°C shall not exceed the values

For items a, b, e, f, i: temperature rise $= 0.9(T_h - T_a)$

For items c, d, g, h: temperature rise $= 0.965(T_h - T_a)$

where T_h, the hot-spot temperature, is given by the following table:

Class	Items *a* and *f*	All other items
A	115°C	105°C
B	140°C	130°C
F	165°C	155°C
H		180°C

Preferred values of ambient temperature above 40°C are 50°C, 65°C, 90°C, 115°C.

The time ratings for single-phase and polyphase induction motors shall be 5, 15, 30, 60 min, and continuous. All short-time ratings are based upon a load test which shall commence when the windings and parts of the motor are within 5°C of the ambient temperature.

Service Factor. General-purpose fractional- and integral-horsepower motors are given a "service factor," which allows the motor to deliver greater than rated horsepower, without damaging its insulation system. The motor is operated at rated voltage and frequency. The standard service factors are 1.4 for motors rated $^1/_{20}$ to $^1/_8$ hp; 1.35 for $^1/_6$ to $^1/_3$ hp; 1.25 for $^1/_2$ hp to the frame size for 1 hp at 3600 r/min. For all larger motors through 200 hp, the service factor is 1.15. For 250 to 500 hp, the service factor is 1.0.

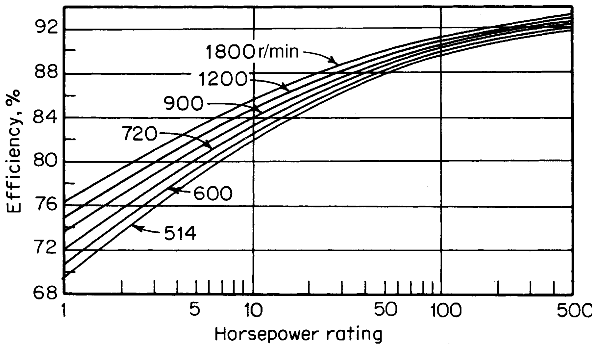


FIGURE 20-36 Typical full-load efficiencies of Design B squirrel-cage motors.

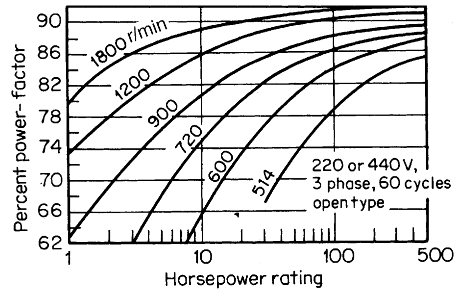


FIGURE 20-37 Typical full-load power factors of Design B squirrel-cage motors.

Efficiency and Power Factor. Typical full-load efficiencies and power factors of standard Design B squirrel cage induction motors are given in Figs. 20-36, and 20-37, respectively. The efficiencies of Design A motors are generally slightly lower, and those of Design D motors considerably lower. The power factors of Design A squirrel-cage induction motors are slightly higher, and those of Design C are slightly lower. Energy-efficient motors are those whose design is optimized to reduce losses. Comparative efficiencies of standard and energy-efficient motors of NEMA Design B are shown in Fig. 20-38.

Full-Load Current. With the efficiency and power factor of a 3-phase motor known, its full-load current may be calculated from the formula

$$\text{Full-load current} = \frac{746 \times \text{hp rating}}{1.73 \times \text{efficiency} \times \text{pf} \times \text{voltage}} \tag{20-29}$$

where the efficiency and power factor are expressed as decimals.

Torques and Starting Currents. Starting and breakdown torques of common Design A, B, and C squirrel-cage induction motors are given in Table 20-6. Relative values for other classes of squirrel-cage

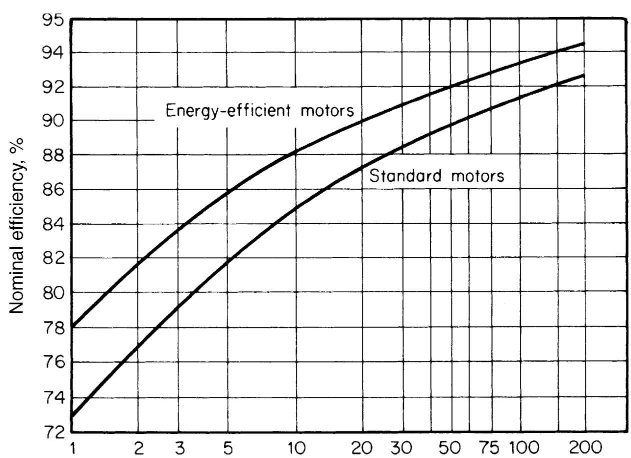


FIGURE 20-38 Nominal efficiencies for NEMA Design B, 4-pole motors, 1800 r/min; standard vs. energy-efficient motors.

TABLE 20-6 Torques-Polyphase Induction Motors
(Percent of full-load torque)

Rpm	3,600		1,800				1,200				900				720	
Torque	LR	BD	LR	LR	BD	BD	LR	LR	BD	BD	LR	LR	BD	BD	LR	BD
Design	AB	B	AB	C	B	C	AB	C	B	C	AB	C	B	C	AB	B
$^1/_2$ hp	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	150	. . .	250	. . .	150	200
$^3/_4$ hp	. . .	. . .	. . .	. . .	. . .	. . .	175	. . .	275	. . .	150	. . .	250	. . .	150	200
1 hp	. . .	. . .	275	. . .	300	. . .	175	. . .	275	. . .	150	. . .	250	. . .	150	200
$1^1/_2$ hp	175	275	265	. . .	300	. . .	175	. . .	275	. . .	150	. . .	250	. . .	150	200
2 hp	175	250	250	. . .	275	. . .	175	. . .	250	. . .	150	. . .	225	. . .	145	200
3 hp	175	250	250	. . .	275	. . .	175	250	250	225	150	225	225	200	135	200
5 hp	150	225	185	250	225	200	160	250	225	200	130	225	225	200	130	200
$7^1/_2$ hp	150	215	175	250	215	190	150	225	215	190	125	200	215	190	120	200
10 hp	150	200	175	250	200	190	140	225	200	190	125	200	200	190	120	200
15 hp	150	200	165	225	200	190	135	200	200	190	125	200	200	190	120	200
20 hp	150	200	150	200	200	190	135	200	200	190	125	200	200	190	120	200
25 hp	150	200	150	200	200	190	135	200	200	190	125	200	200	190	120	200
30 hp	150	200	150	200	200	190	135	200	200	190	125	200	200	190	120	200
40–200 hp	—*	200	—*	200	200	190	—*	200	200	190	125	200	200	190	120	200

*Progressively lower values for these larger ratings.
Note: LR = locked-rotor torque; BD = breakdown torque; A, B, and C refer to Design A, etc.

TABLE 20-7 Locked-Rotor kVA for Code-Letter Motors

Code letter*	kVa per hp, with locked rotor	Codek letter*	kVa per hp with locked rotor
A	0–3.14	L	9.0–9.99
B	3.15–3.54	M	10.0–11.19
C	3.55–3.99	N	11.2–12.49
D	4.0–4.49	P	12.5–13.99
E	4.5–4.99	R	14.0–15.99
F	5.0–5.59	S	16.0–17.99
G	5.6–6.29	T	18.0–19.99
H	6.3–7.09	U	20.0–22.39
J	7.1–7.99	V	22.4 and up
K	8.0–8.99		

National Electrical Code.

motors are indicated by the curves of Fig. 20-33. The minimum breakdown torque for wound-rotor motors is 200% of full-load torque. As indicated by the curves of Fig. 20-34, the starting torque and starting current of wound-rotor motors vary with the amount of external resistance in the secondary circuit.

The starting kVA of a squirrel-cage motor is indicated by a code letter stamped on the nameplate. Table 20-7 gives the corresponding kVA for each code letter, and the locked-rotor current can be determined from

$$\text{Locked-rotor current} = \frac{\text{kVA/hp} \times \text{hp} \times 1000}{k \times \text{line volts}} \tag{20-30}$$

where $k = 1$ for single-phase, and $k = 1.73$ for 3-phase.

Maximum locked-rotor current for Design B, C, and D 3-phase motors has been standardized as shown in Table 20-8 for 230 V. The starting current for motors designed for other voltages is inversely proportional to the voltage.

Starting Methods. Wound-rotor motors are invariably started on full voltage but with external resistance in the secondary circuit. Ordinarily sufficient resistance is provided to give 100% torque at standstill, which means that 100% current will be drawn from the line. If a higher torque is required to start the load, less external resistance must be used, and the current drawn is proportionately higher. As the motor accelerates, the external secondary resistance is short-circuited in one or more steps.

The locked-rotor values in Table 20-8 are generally recognized as the minimum needed by motor designers to obtain the required torque characteristics for general-purpose motors. Squirrel-cage motors with these values are usually acceptable for full-voltage starting on power lines and also on combined light and power secondaries of 208 or 230 V, if manually controlled (infrequently started).

TABLE 20-8 Locked-Rotor Current for 3-Phase Motors at 230 V

Rated horsepower	Classes B, C, D, amperes	Rated horsepower	Classes B, C, D, amperes	Rated horsepower	Classes B, C, D, amperes	Rated horsepower	Classes B, C, D, amperes	Rated horsepower	Class B amperes
1	30	$7^1/_2$	127	30	435	100	1450	250	3650
$1^1/_2$	40	10	162	40	580	125	1815	300	4400
2	50	15	232	50	725	150	2170	350	5100
3	64	20	290	60	870	200	2900	400	5800
5	92	25	365	75	1085			450	6500
								500	7250

In the case of automatically controlled (frequently started) equipment, with 208- or 230-V motors supplied from combined light and power secondaries, current-reducing starters to reduce the current to about 65% of these values may be required, unless consultation with the power company indicates that the available system capacity will permit use of full-voltage starting. In any case, consultations with the power company for motor applications above 25 hp are advisable.

Autotransformer starters (compensators) are the most popular of any reduced-voltage type. They have the advantage that the ratio of torque developed by the motor to the current drawn from the line remains substantially the same as for full-voltage starting. The motor torque and the current drawn from the line (neglecting the magnetizing current of the autotransformer) are both reduced in proportion to the square of the voltage impressed on the motor. The magnetizing current of the autotransformer generally does not exceed 25% of motor full-load current. Normally, the motor accelerates nearly to full speed on the reduced-voltage connection and is then transferred to full voltage. Since the circuit to the motor is opened and then immediately reclosed, a transient inrush of current occurs which may be of much greater magnitude than the current normally drawn by the motor at the speed at which the transfer is made. This transient inrush, however, is of such extremely short duration that it does not produce an objectionable voltage disturbance on the average power system. Standard autotransformer starters are provided with 65% and 80% voltage taps in sizes up to 50 hp and with 50%, 65%, and 80% voltage taps in the larger sizes.

"Part-winding" starting is being more widely used for reducing starting current. This involves arranging the stator winding so that, by use of adequate control devices, one part of the stator winding is first energized and subsequently the remainder of the winding is energized in one or more steps. The purpose is to reduce the initial values of the starting current drawn and/or the starting torque developed by the motor. The usual arrangement involves energizing one-half the stator winding on the first step, resulting in approximately 50% of normal locked-rotor torque and approximately 60% of normal locked-rotor current. While this torque may be insufficient to start the motor in some applications, it permits drawing full-winding starting current from the system in two increments. Another method is to connect two-thirds of the winding on the first step, by using a 4-pole contactor, in which case the motor should accelerate promptly to full speed. The remaining third of the winding is then connected by closing a second contactor with only two poles.[2–4]

Resistor-type reduced-voltage starters are sometimes used. They have the disadvantage that the current drawn from the line is reduced in direct ratio to the impressed voltage, while the torque developed by the motor is reduced as the square of this voltage. The resistor is short-circuited, either all at once or in steps, when the motor comes up to speed. The circuit for the motor is not broken in transferring to full voltage, as is the case with the autotransformer starter. These features make the resistor-type starter adapted for use where "increment-type" starting-current restrictions exist. With the resistor-type starter, the contactors, which short-circuit the resistors as well as the line contactors, must carry the full current of the motor, whereas in part-winding starting, the contactors for the two parts of the winding each carry only half the total current.

Reactor-type reduced-voltage starters are sometimes used on larger motors, most frequently on high-voltage motors (2300 V or above), where oil circuit breakers are necessary to provide sufficient current-interrupting capacity. In such cases, the reactor and starting circuit breaker are placed in the neutral of the motor. The breaker can then be of low-interrupting capacity, since the fault current at this point is limited by the reactance of the motor windings.

Wye-delta starting, though quite common abroad, is used in the United States primarily for refrigeration compressors. This starter consists of a switching arrangement that transfers the motor winding from Y for starting to Δ for running. The current drawn and the torque developed by the motor are thus reduced to only one-third their full voltage values. This very low torque, the extra contactors required, and the current inrush when the circuit is reclosed on Δ make this scheme less attractive than others.

Motors are frequently supplied from power systems consisting of complex networks for which calculation of the voltage drop would be difficult. The voltage drop may be estimated, however, if the short-circuit kVA is known at the point of power delivery.

When motor-starting kVA is drawn from a system, the voltage drop in percent of the initial voltage is approximately equal to 100 times the motor-starting kVA divided by the sum of this kVA and the short-circuit kVA. The motor-starting kVA used should be that drawn by the motor if the initial system voltage is maintained. For example, if a 1000-hp motor has a starting kVA of 5000 at the initial system voltage and the system short-circuit kVA is 50,000, the voltage drop will be approximately

$$\frac{5000 \times 100}{(5000 + 50{,}000)} = 9\% \text{ of initial voltage}$$

In many systems, the short-circuit kVA varies over a wide range depending on the number of parallel lines in service, service interconnections, etc. While the highest short-circuit kVA is of interest for circuit interruption, the minimum short-circuit kVA should be used for voltage-drop calculations since it gives the highest value.

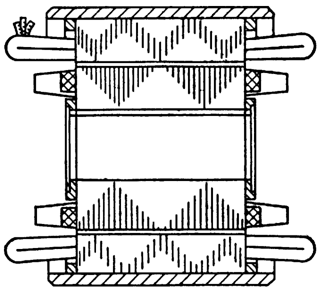


FIGURE 20-39 Cross section of shell-type motor.

Shell-Type Motors. These motors consist of stators and rotors only, without shafts, end shields, bearings, or conventional frame (Fig. 20-39). The rotors are mounted directly on a shaft of the driven machine, which must also include a suitable support for the stator and a ventilating arrangement. The motors are built with relatively small outside diameters but may be slightly longer than standard machines. Furthermore, horsepower ratings over a rather wide range are built in each frame diameter, the ratings for the different diameters overlapping slightly. Although a great many of the motors used are for operation at standard commercial frequencies giving speeds up to 3600 r/min (on 60 Hz), they are frequently supplied for operation at higher frequencies and correspondingly higher speeds. Frequencies up to 2000 Hz with a corresponding 2-pole motor speed of 120,000 r/min have been used, but the more common "high" frequencies range from 60 to 240 Hz, giving 2-pole motor speeds up to 14,400 r/min.

Shell-type motors are used principally on machine tools and woodworking machinery. Their relatively small physical size facilitates a compact design with maximum flexibility in arrangement of machine parts. The small diameter of the motors is of particular value, since it allows close spacing of spindle shafts.

The wide range of ratings available in each diameter reduces the cost of providing suitable mountings for the motors.

Motors of similar mechanical construction but with special insulation are used in hermetically sealed refrigeration and air-conditioning compressors, where the motor runs in an atmosphere of refrigerating gas. The insulation must neither harm nor be harmed by the refrigerant and, so that the refrigerant may be kept clean and dry, must not trap moisture or dirt.

Dimensions. NEMA has standardized mounting dimensions for various types of motors, those standardized for polyphase induction motors covering ratings from 1 to 125 hp (at 1800 r/min). For convenience each set of standardized dimensions has been assigned a frame number, and the various ratings of motors have been assigned frame numbers from the series. Any motor offered by a manufacturer having a frame number from this series will have the corresponding standardized mounting dimensions. These are listed in *NEMA Motor and Generator Standard*, Publ. MG1-1987.

References on Polyphase Induction Motors

1. NEMA, *Motor and General Standards*, Publ. MG1-1987.
2. Alger, P. L., Ward, H. C., Jr., and Wright, F. H., Split-Winding Starting in 3-Phase Motors, *Trans. AIEE*, 1951, vol. 70, pt. 1, p. 867.
3. Alger, P. L., and Agacinsky, L., A New Method for Part-Winding Starting of Polyphase Motors, *Trans. AIEE*, *Power Apparatus and Systems*, Feb. 1956, no. 22, p. 1455.
4. Alger, P. L., Performance Calculations for Part-Winding Starting of Three-Phase Motors, AIEE Conf. Paper, pp. 56–515.

20.4.4 Single-Phase Induction Motors

General Theory. If one supply line to a polyphase induction motor is opened, the motor will not develop any starting torque, although if it is already operating, it will continue to run at slightly reduced speed, with a somewhat lower breakdown torque. The crux of the single-phase motor problem, therefore, is in providing auxiliary means for starting.

The magnetic field of a single-phase winding carrying alternating current may be represented as a phasor stationary in space but alternating in time, or as the sum of two equal and oppositely revolving field phasors, which are constant in magnitude. In a polyphase motor, the backward-revolving field phasors of the several phases cancel each other, and the forward-revolving ones add directly, giving a uniform revolving field. In the single-phase motor, means are provided to reduce the backward field, but this field has always some remaining magnitude (except at one particular load in the case of certain capacitor-run motors), and consequently a single-phase induction motor always has extra losses and a double-frequency pulsating torque not possessed by a polyphase motor.

A simple way to visualize the effects of this backward field is to consider that the forward- and backward-revolving fields are separately produced by the same stator current; that is, they are connected in series. Each field may then be treated as a separate polyphase induction motor, the forward field having a slip s with respect to the rotor, and the other a slip $2 - s$. At standstill, both values of slip are unity, and the two circuits are identical. At all times, the net torque developed is equal to the difference of the separate torques produced by the two fields. On this basis, the single-phase induction motor equivalent circuit is given by Fig. 20-40.

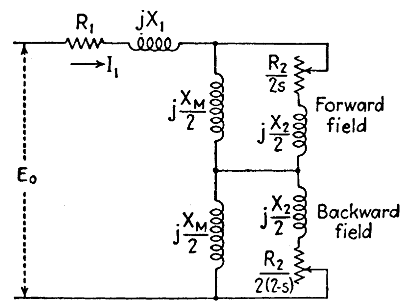


FIGURE 20-40 Equivalent circuit of single-phase induction motor.

The values of R_1, X_1, R_2, X_2, and X_M are the impedance constants derived by measurements across the single-phase terminals. Since half the total air-gap impedance at standstill is due to each field, the magnetizing and secondary impedance values are divided by 2 to obtain the values corresponding to the separate fields.

Inspection of this circuit reveals several interesting properties of the motor. At full speed, s is very small, and the backward field appears as an external series impedance of $R_2/4 + j(X_2/2)$. The corresponding loss $I^2R_2/4$ represents the power delivered to the rotor by the backward field. However, there is an equal loss due to the rotor's being driven forward at speed $1 - s$ against the backward-field torque; so the total loss caused by the backward field is $I^2R_2/2$, approximately. Since the backward-field rotor currents occur at double-line frequency, any double squirrel-cage or deep-bar rotor design which had an increased resistance at high frequency would greatly increase the power losses, and such designs, therefore, are seldom used for single-phase motors. The breakdown torque of a single-phase motor may be approximately calculated for a polyphase induction motor, if the impedance of the backward-revolving field is considered as a series impedance added in the primary circuit of the polyphase motor. Hence, any increase in the secondary resistance of a single-phase motor actually reduces the breakdown torque and lowers the speed at which breakdown occurs.

Another interesting characteristic is the double-frequency torque pulsation. The double-frequency current in the rotor reacting on the slip-frequency forward magnetic field evidently produces a torque pulsation, even at no load. Physically, the no-load part of the pulsating torque provides the means for supplying and removing the magnetic field twice each cycle in the axis at right angles to the stator winding, and the additional part under load corresponds to the double-frequency pulsation of the single-phase power input to the rotor. To prevent objectionable transmitted vibration and noise from this cause, it is usual to mount single-phase machines on supports with torsional elasticity of some type, often rubber rings encircling the bearing housings in the case of fractional-horsepower motors.

Shaded-Pole Motor. The simplest way of providing a single-phase induction motor with starting torque is to place a permanently short-circuited winding of relatively high resistance in the stator at an electrical angle of 30° to 60° from the main winding. Usually this auxiliary winding, called a "shading coil," consists of an uninsulated copper strip encircling approximately one-third of a pole pitch. The current induced in the shading coil, by the portion of the main field linking it, reduces the magnitude of this flux and also causes it to lag in time phase. In consequence, the air-gap field has two components, an undamped alternating flux and a damped flux displaced both in space and in time. Shaded-pole motors are used only in very small sizes normally below 50 W output. Principal applications are for desk fans and air circulators, where their simplicity, low torque, and low cost are well suited to the requirements.

The inherently high slip of a shaded-pole motor makes it convenient to obtain speed variation on a fan load by reducing the impressed voltage. It is common practice to provide multispeed fan operation by employing a small switched-series reactor or a phase-controlled solid-state device to control the motor voltage.

Resistance Split-Phase Motors. A considerably greater starting torque can be obtained by providing a separate starting winding, or auxiliary phase, 90° displaced in space from the main winding of a single-phase induction motor. This extra winding is normally wound with fewer turns of a much smaller size of wire, so that it has a considerably greater resistance to reactance ratio than the main winding, and it is connected directly across the power supply, in parallel with the main winding. Just as in the case of the shaded-pole motor, the field of the auxiliary winding is displaced in time and in space, so that its vectorial combination with the main field gives a much larger forward than backward field component. The motor can be reversed by reversing either the main or the auxiliary winding.

Since the auxiliary winding is normally located 90° from the main winding, the two are mutually noninductive at standstill and the standstill characteristics may be calculated from two independent circuits each like that of Fig. 20-40. By a similar analysis to that of the preceding section, the starting torque of a split-phase motor is

$$T = \frac{14.1aK}{N_s} I_M I_A R_2 \sin\theta \quad \text{ft} \cdot \text{lb} \tag{20-31}$$

where I_M = main winding starting current in amperes; I_A = auxiliary winding starting current in amperes; N_s = synchronous speed; R_2 = resistance component of standstill impedance of main winding, less the primary resistance; a = ratio of effective number of turns in auxiliary winding to main-winding effective turns; K = an empirical coefficient, which allows for nonfundamental rotor losses, usually equal to 0.9; and θ = angle of phase split between I_M and I_A.

Design limitations usually prevent θ from being greater than 30°, so that the starting torque per voltampere cannot exceed half that of a 2-phase motor built in the same parts. Since, in addition, both I_M and I_A are drawn from a single phase of the power supply, the starting current is excessive, limiting the use of the resistance split-phase motor to sizes below $^1/_3$ hp.

The auxiliary winding is opened automatically as the motor approaches full speed, as otherwise prohibitive losses would occur in it. Usually this is accomplished by means of a centrifugal switch or, in the case of hermetically sealed motors, by an electromagnetic relay. The high current density used to obtain an adequate resistance value makes the initial rate of temperature rise of the auxiliary winding very great, sometimes more than 50°C/s, so that these motors are not satisfactory for repeated starting or for inertia loads.

Earlier split-phase motor designs included motors with stationary external squirrel-cage members, with the primary windings on the rotor, receiving their power through slip rings. They are now normally built, however, with uniformly distributed partly closed stator slots, enameled-wire concentric stator windings, and a cast-aluminum or welded-copper squirrel cage on the rotor.

Typical characteristic curves for a $^1/_6$-hp 60-Hz 1725-r/min, resistance split-phase motor are shown in Table 20-11.

Repulsion-Start Induction-Run Motor. A common way of obtaining single-phase-motor starting torque is to provide a dc winding and commutator on the rotor, with a single pair of short-circuited brushes for starting and a centrifugal mechanism, which short-circuits the entire commutator as the motor approaches full speed. This gives a pure repulsion-motor starting characteristic with very high torque per ampere and pure single-phase induction-motor operating characteristics. These motors are widely used in sizes up to about 5 hp. Typical characteristics of a 1-hp 60-Hz 1800-r/min motor of this type are shown in Table 20-11.

Capacitor Motors. Low-cost low-voltage capacitors have proved extremely useful in improving the performance of split-phase motors. By inserting an external series capacitor in the auxiliary winding circuit and making this winding with many more turns of much lower resistance, the angle of phase split θ can be increased to 90°, or even more, and the coincident increase in the turn ratio a permits a further decrease in the auxiliary winding current. Thus, the capacitor-start motor gives an adequate starting torque for a reasonable starting current and at the same time has so much greater thermal capacity than a resistance split-phase motor, by virtue of the reduced winding-current density, that it is satisfactory for nearly all industrial single-phase motor applications.

Figure 20-41 illustrates a convenient method of determining the best size of capacitor to use with a given motor. I_M represents the locked-rotor current in the main winding and I_A the current in the auxiliary winding. With no external capacitor, $X_C = 0$, and the motor becomes a plain resistance split type. As X_C is increased, I_A moves ahead in time phase, following a circular locus, increasing the torque and reducing the total current drawn from the line. Points of maximum starting torque and maximum starting torque per ampere are indicated on the diagram.

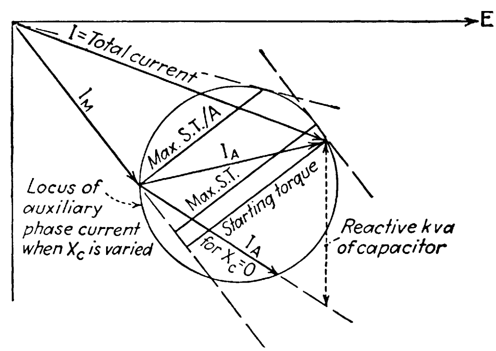


FIGURE 20-41 Capacitor-motor, starting-torque diagram.

Usually low-voltage electrolytic capacitors are used for starting purposes, since these are economical in 115-V intermittent ratings. However, in cases of severe starting duty, higher-voltage motors, or where capacitors are retained in the circuit during operation, paper or film-type, oil-filled ac capacitors are used.

For most applications, the auxiliary winding is opened by a centrifugal switch or relay, as the motor approaches full speed, just as in the case of the resistance split-phase motor. Such motors are called *capacitor-start motors*. In some cases of smaller-sized motors with low-starting-torque requirements, however, it is permissible to leave the capacitor permanently in circuit. These are called *permanent-split capacitor motors*. The limitations of starting torque and motor size on this type are the result of the inherent tendency of the auxiliary-winding current to increase in magnitude and shift backward in time phase as the motor accelerates, so that unless the capacitor impedance is very high, the motor will have objectionable losses and large torque pulsations at full speed. However, the power loss in the capacitor circuit at speed is very much less than in a shading coil for a given starting torque, and so the permanent-split capacitor motor is finding increasing use for fan drive in sizes up to $^1/_4$ hp.

For the larger capacitor motors, in sizes of $^1/_2$ hp and up, it is frequently economical to retain the auxiliary winding in circuits with a reduced capacitor size, to improve the operating characteristics. This is usually accomplished by providing a large electrolytic or highly stressed capacitor in parallel with a small oil capacitor at starting and cutting the former out of circuits with a centrifugal switch or relay when the motor approaches full speed. Such motors are called *capacitor-run motors* and have winding connections as shown in Fig. 20-42. The analysis of the capacitor motor

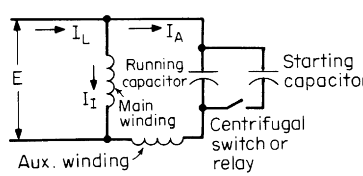


FIGURE 20-42 Capacitor-run, motor-winding connections.

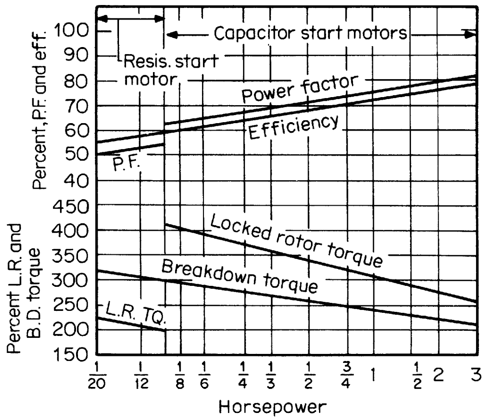


FIGURE 20-43 Typical operating characteristics of 1800-r/min single-phase motors.

TABLE 20-9 Standard Horsepower and Speed Ratings—Single-Phase Constant-Speed Motors

Standard horsepower ratings					
$^1/_{20}$	$^1/_6$	$^1/_2$	$1^1/_2$	5	15
$^1/_{12}$	$^1/_4$	$^3/_4$	2	$7^1/_2$	20
$^1/_8$	$^1/_3$	1	3	10	25

Standard speed ratings		
Rpm 60 cycles	Fractional hp	Integral hp
3600	$^1/_{20}-1$	$1^1/_2-25$
1800	$^1/_{20}-^3/_4$	$1-25$
1200	$^1/_{20}-^1/_2$	$^3/_4-25$
900	$^1/_{20}-^1/_3$	$^1/_2-25$
Rpm 50 cycles		
3000	$^1/_{20}-1$	$1^1/_2-20$
1500	$^1/_{20}-^3/_4$	$1-20$
1000	$^1/_{20}-^1/_2$	$^3/_4-20$
750	—	$^1/_2-20$

is done with an equivalent circuit of the type shown in Fig. 20-43 with the capacitors properly introduced.

Horsepower, Speed, and Voltage Ratings. Standard horsepower and speed ratings of single-phase motors are given in Table 20-9. Motors built in frames having a continuous rating of less than 1 hp, open type, at 1700 to 1800 r/min are designated *fractional-horsepower* motors, and those built in larger frames are called *integral-horsepower* motors. Somewhat different standards of performance have been established for the two classes. ANSI C-50 and *NEMA Motor and Generator Standards*, Publ. MG1-1978, include basic standards for both fractional-and integral-horsepower motors which are normally followed in specifications and testing.

Both capacitor and split-phase motors are available in the multispeed as well as the single-speed type. They are used principally for belt and direct drive of centrifugal and propeller fans and are of the variable-torque class. The multispeed motors for fan drive allow a change in fan speed without changing pulleys, which is essential where remote or automatic control of the rate of air delivery is required.

The standard voltage ratings for single-phase motors are 115 and 230 V for supply lines rated 120 and 240 V. Power companies place a limit on the size of motors that may be connected to single-phase lines. The limit usually falls between $^1/_2$ and 1 hp for 120-V circuits and between 3 and 10 hp for 240-V circuits.

Temperature Rise. The standard temperature rises and service factors for single-phase motors are the same as for polyphase motors.

Efficiencies and Power Factors. Typical efficiencies and power factors of the various types of induction motor that might be used to fill the requirements of the different ratings are shown in Fig. 20-43. Repulsion-start induction-run motors have about the same efficiencies and power factors except in the $1^1/_2$-to 3-hp range, where they are lower. Repulsion-induction motors have roughly the same efficiencies but higher power factors.

Single-Phase Motor Characteristics. The full-load current of a single-phase motor is equal to

$$\frac{746 \times \text{hp}}{\text{Efficiency} \times \text{voltage}} \qquad (20\text{-}32)$$

where the efficiency and power factor are expressed as decimals. Approximate values of full-load current are given in Table 20-10. These are used for selecting wire and fuse sizes if no more accurate data are available.

TABLE 20-10 Single-Phase Motor Characteristics

	Hp	Approximate full load, A		Locked rotor, A				Breakdown torque (for defining hp ratings), oz · ft above line; lb · ft below line			
		115 V	230 V	115 V		230 V		3600 rpm	1800 rpm	1200 rpm	900 rpm
Fractional hp	1/6	4.4	2.2	20		10		8.7–11.5	16.5–21.5	24.1–31.5	31.5–40.5
	1/4	5.8	2.9	23		11 1/2		11.5–16.5	21.5–31.5	31.5–44.0	40.5–58.0
	1/3	7.2	3.6	31		15 1/2		16.5–21.5	31.5–40.5	44.0–58.0	58.0–77.0
	1/2	9.8	4.9	45		22 1/2		21.5–31.5	40.5–58.0	58.0–82.5	. . . . . .
	3/4	13.8	6.9	61		30 1/2		31.5–44.0	58.0–82.5	5.16–6.9	. . . . . .
				Design		Design					
				L	M	L	M				
Integral hp	1	16	8	70	. . .	35	. . .	44.0–58.0	5.16–6.8	6.9–9.2	
	1 1/2	20	10	. . .	. . .	50	40	3.6–4.6	6.8–10.1	9.2–13.8	
	2	24	12	. . .	. . .	65	50	4.6–6.0	10.1–13.0	13.8–18.0	
	3	34	17	. . .	. . .	90	70	6.0–8.6	13.0–19.0	18.0–25.8	
	5	56	28	. . .	. . .	135	100	8.6–13.5	19.0–30.0	25.8–40.5	
	7 1/2	80	40	. . .	. . .	200	150	13.5–20.0	30.0–45.0	40.5–60.0	
	10	100	50	. . .	. . .	260	200	20.0–27.0	45.0–60.0	. . . . . .	

Note: 1 oz · in = 0.00706 N · m; 1 lb · ft = 1.356 N · m.

Characteristics of a 60-Hz, 4-pole, 1800-r/min, single-phase motor are shown in Fig. 20-43.

The horsepower rating of a single-phase motor is defined by its breakdown torque. Thus, any 1800-r/min motor with a breakdown torque between 31.5 and 40.5 oz · ft is, by definition, a $^1/_3$-hp motor. The value used for definition is the minimum of the range of manufacturing variation for that particular design.

Starting Current. Maximum values of locked-rotor current are established by NEMA for 60-Hz motors as shown in Table 20-10. In integral-horsepower sizes, NEMA has established two sets of locked-rotor values. The Design L motors include those types having inherently higher locked-rotor values. The Design L motors include those types having inherently higher locked-rotor current than Design M motors.

20.5 OTHER TYPES OF ELECTRIC MOTORS AND RELATED APPARATUS

Induction Generators. Any induction motor, if driven above its synchronous speed when connected to an ac power source, will deliver power to the external circuit. The generator operation is easily visualized from the equivalent circuit of Fig. 20-28, corresponding to negative slip. The induction generator must always take reactive power from the load or the line for excitation and for the I^2X losses. For this reason, the induction generator can only operate in parallel with an electric power system or independently with a load supplemented by capacitors. For independent operation, the speed must be increased with load to maintain constant frequency; the voltage is controlled with the capacitors.

An induction generator delivers an instantaneous 3-phase short-circuit current equal to the terminal voltage divided by its locked-rotor impedance. Its rate of decay is much faster than that of a synchronous generator of the same rating, corresponding to the subtransient time constant T'_{do}; sustained short-circuit current is zero.

The virtue of the induction generator is its ability to self-synchronize when the stator circuit is closed to a power system. At one time induction generators were used for small, unattended hydro stations. Since the late 1990s, induction generators have been used in a similar manner for wind turbines and cogeneration units. They have also been used for high-speed, high-frequency generators, because of their squirrel-cage rotor construction.

Synchronous Induction Motors. There are three types of motors that can start and run as induction motors yet can lock into the supply frequency and run as synchronous motors as well. They are (1) the wound-rotor motor with dc exciter, (2) the permanent-magnet (PM) synchronous motor, and (3) the reluctance-synchronous motor. The latter two types have been used primarily with adjustable-frequency inverter power supplies. In Europe, wound-rotor induction motors have often been provided with low-voltage dc exciters that supply direct current to the rotor, making them operate as synchronous machines. With secondary rheostats for starting, such a motor gives the low starting current and high torque of the wound-rotor induction motor and an improved power factor under load. Several different forms of these synchronous induction motors have been proposed, but they have not shown any net advantage over usual salient-pole synchronous or induction machines and are very seldom used in the United States. The PM synchronous motor is shown in Fig. 20-44*a*. The construction is the same as that of an ordinary squirrel-cage motor (either single or polyphase), except that the depth of rotor core below the squirrel-cage bars is very shallow, just enough to carry the rotor flux under locked-rotor conditions. Inside this shallow rotor core is placed a permanent magnet, fully magnetized. The rotor core serves as a keeper, so that the rotor is not demagnetized by removing it from the stator. In starting, the rotor flux is confined to the laminated core. As the speed rises, the rotor frequency decreases and the rotor flux builds up, creating a pulsating torque with the field of the magnet, as when a synchronous motor is being synchronized after the dc field has been applied. As the motor approaches full speed, therefore, the ac impressed field locks into step with the field of the magnet and the machine runs as a synchronous motor. The absence of rotor I^2R loss,

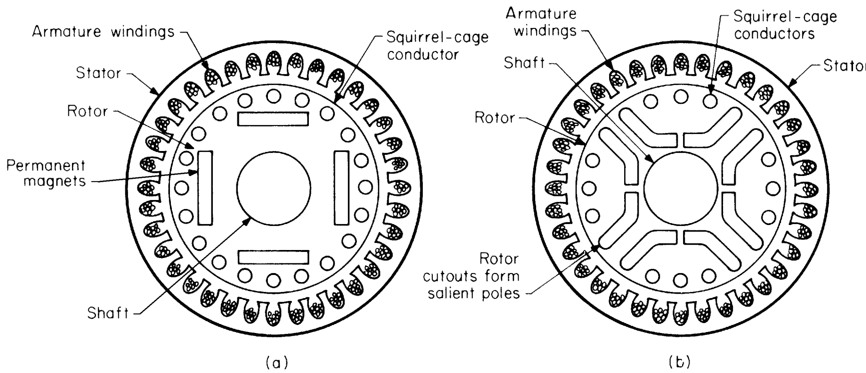


FIGURE 20-44 Cross section of (*a*) a conventional PM synchronous motor and (*b*) a reluctance-synchronous motor.

the synchronous speed operation, and the high efficiency and power factor make the motor very attractive for special applications, such as high-frequency spinning motors. When many such motors are supplied from a high-frequency source, the kVA requirements are reduced to perhaps 50% of those needed for usual induction motor types, with consequent large savings.

If the rotor surface of a *P*-pole squirrel-cage motor is cut away at symmetrically spaced points, forming *P* salient poles, the motor will accelerate to full speed as an induction motor and then lock into step and operate as a synchronous motor. The synchronizing torque is due to the change in reluctance and, therefore, in stored magnetic energy, when the air-gap flux moves from the low- into the high-reluctance region. Such motors are often used in small-horsepower sizes, when synchronous operation is required, but they have inherently low pull-out torque and low power factor, and also poor efficiency, and therefore require larger frames than the same horsepower induction motor. The PM synchronous motor has superior performance in every way, except possibly cost. A cross section of the reluctance-synchronous motor is shown in Fig. 20-44*b*. These motors are available up to about 5 hp.

If the number of rotor salients is nP, instead of P, and if the P-pole motor winding is arranged to also produce a field of $(n - 1)P$ or $(n - 1)P$ poles, the motor may lock into step at a subsynchronous speed and run as a subsynchronous motor. For the P-pole fundamental mmf, acting on the varying rotor permeance will create $(n + 1)P$ and $(n - 1)$ P-pole fields from this case, and these will lock into step with the independently produced $(n - 1)P$- or $(n + 1)P$-pole field, when the rotor speed is such as to make the two harmonic fields turn at the same speed in the same direction.

It is difficult to provide much torque in such subsynchronous motors, and their use is therefore limited to very small sizes, such as may be used in small timer or instrument motors.

Linear Motors. Linear induction motors (LIMs) have been built in fractional-horse-power ratings for such applications as moving drapes, and up to several thousand horsepower for driving tracked air-cushion transit vehicles on a guideway. Other applications include moving freight cars in yards, driving people-mover vehicles, and providing reciprocating motion for machine tools. LIMs are built like rotary induction motors with distributed multipole polyphase windings placed in the slots of a plane laminated status as shown in Fig. 20-45. When the windings are excited by a polyphase voltage of frequency f, an air-gap space flux wave is propagated along the length of the stator at a velocity of $v = 2fp$, where p is the pole pitch. The rotor consists of an aluminum or copper sheet, which is propelled by the field with a slip velocity to provide the required thrust. LIMs are either double-sided, with two facing stators operating on a single rotor, or single-sided, with the rotor sheet backed by a moving or stationary magnetic return path. The magnetic force density normal to the stator surface is considerable compared to the tangential force density that moves the rotor, which requires that the stator be well braced mechanically to maintain constant air-gap distances over the surface of the stator. The typical tangential force density is about 3 lb/in^2 for air-cooled windings, where the normal force density is about 30 lb/in^2.

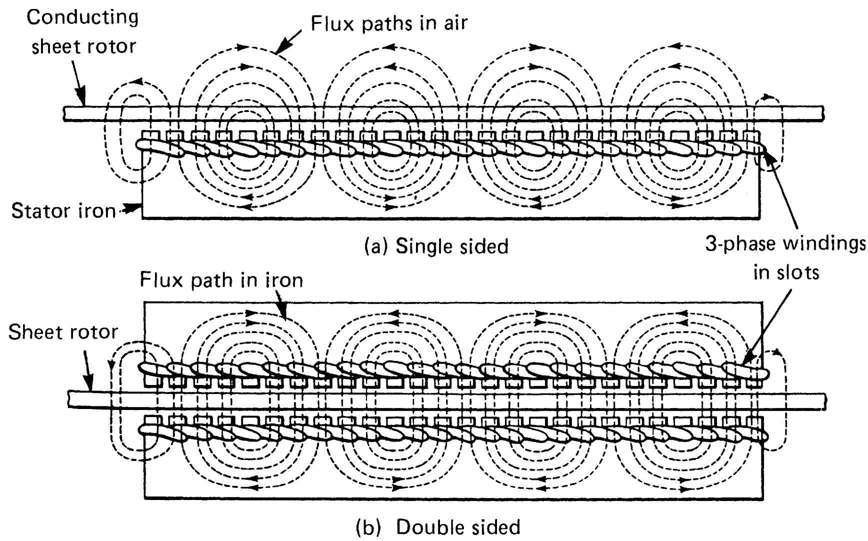


FIGURE 20-45 (*a*) Single-sided and (*b*) double-sided linear induction motors (LIMs) with sheet rotor.

The magnetic air gap of a double-sided LIM is the thickness of the sheet rotor plus the clearance between the rotor and the stators on either side. Whereas most rotary induction motors are built with an air gap of 0.025 to 0.1 mils, the air gap in the LIM is 0.25 to 1.5 in. For this reason, the magnetizing reactance of the LIM is lower than that of an equivalent rotary induction motor. Also the stator leakage reactance is higher. The equivalent circuit of the LIM is shown in Fig. 20-46*a*. Figure 20-46*b* shows the thrust-slip power factor and efficiency curves of a double-sided LIM. This LIM has an air gap of 1.47 in, a rotor sheet thickness of 0.25 in, and a stator length of 9.8 in. The two 3-phase windings of the stator are excited at 173 Hz from inverters to produce a linear synchronous velocity of 395 ft/s. Speed control and breaking of LIMs is done in the same way as in the rotary induction motors.

High-Frequency Motors. For high-speed tools and for spinning of rayon and other threads, a variety of interesting motor constructions have been developed. Normally these are 2-pole 3-phase motors, with special high-frequency power supply of 90, 120, or 180 Hz, giving operating speeds between 500 and 10,500 r/min and up to 25 hp. In textile applications, the motors usually drive individual spinning buckets, which are subject to considerable unbalance due to uneven building up of

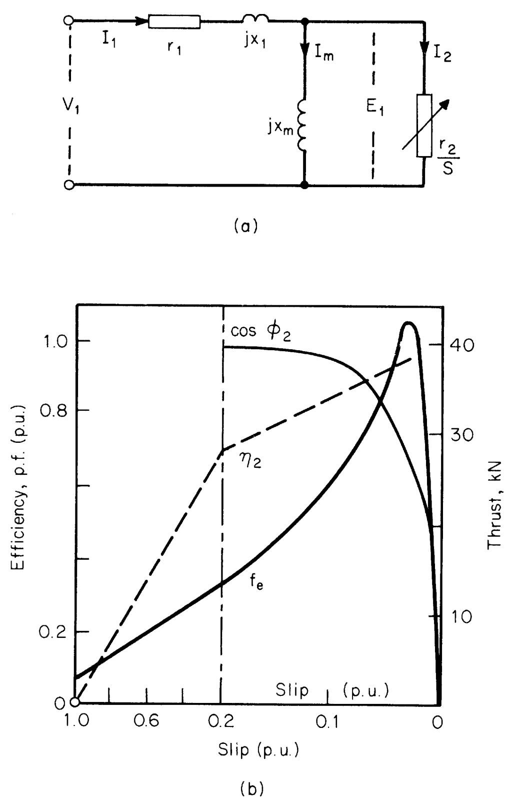


FIGURE 20-46 (*a*) Equivalent circuit of a double-sided LIM, (*b*) the characteristic curves of a typical LIM.

thread, etc. The continual starting and stopping for loading and unloading the buckets requires the motors to carry unbalance reliability through the entire speed range, necessitating careful design of mounting flexibility and shaft stiffness. Most usual applications, however, are in woodworking and similar industries, where separate motor stators and rotors are supplied to the tool manufacturers for building into their particular devices. These motors were powered from high-frequency alternators, but are now powered by adjustable-frequency solid-state inverters.

Three-phase 400-Hz power systems, used on large airplanes, have led to the development of 400-Hz motors with speeds of 12,000 and 24,000 r/min, having weights averaging 2 lb/hp for motors of 1 to 15 hp with 5-min ratings. These motors are open, with an external fan to force air over the windings.

Stepper Motors. The primary characteristic of a stepper motor is its ability to rotate a prescribed small angle (step) in response to each control pulse applied to its windings. Below about 200 pulses per second, the motor rotates in discrete steps in synchrony with the pulses; at higher frequencies up to 16,000 pulses per second, the motor skews without stopping between pulses. Although motors are available for step angles of 90° to 0.180°, the common step is 1.8°. Stepper motors are categorized as permanent-magnet rotor (PM), variable reluctance (VR), or hybrid (PM-VR). The rotor of the PM aligns itself with the energized stator poles as shown in Fig. 20-47*b*. The rotor turns until

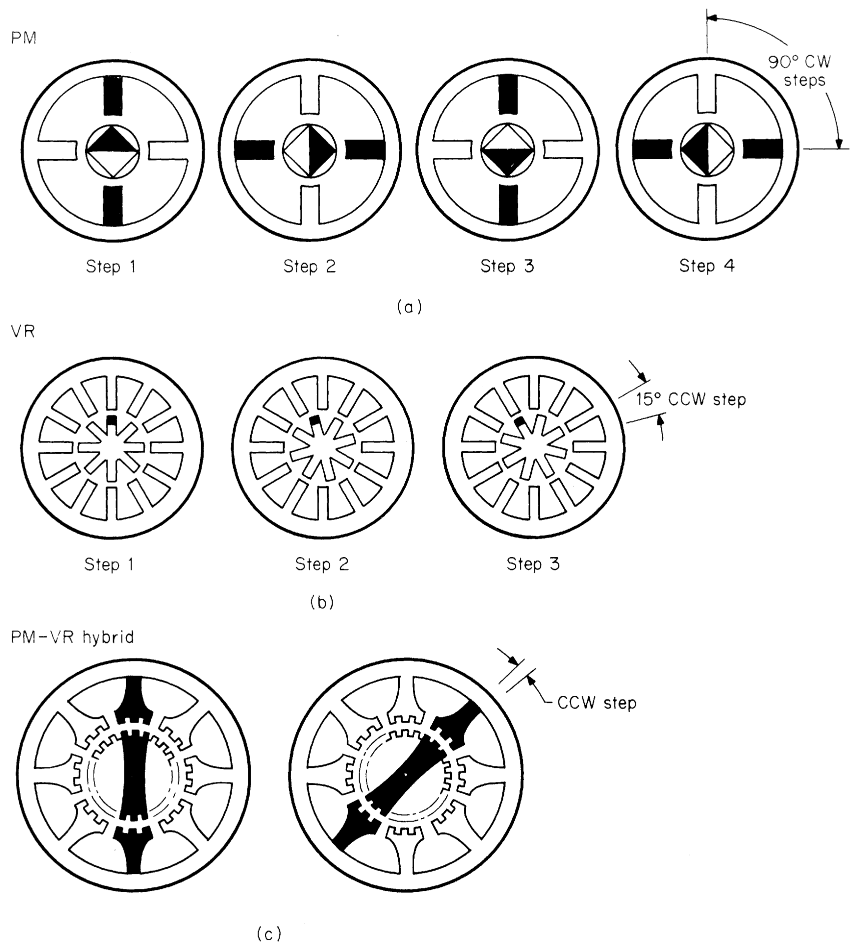


FIGURE 20-47 Three types of stepper motor.

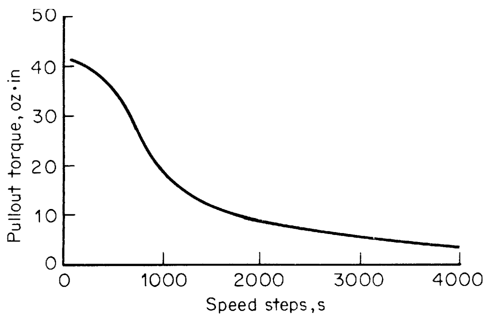


FIGURE 20-48 Pull-out torque vs. speed for a 4-phase 5° step VR step motor running at half steps ($2^1/_2$°).

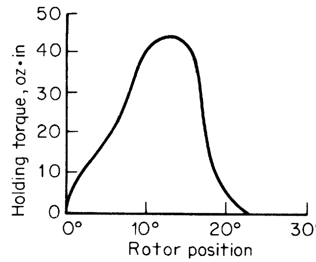


FIGURE 20-49 Holding torque vs. rotor position for a 3-phase, VR step motor, 24 steps per revolution, bidirectional, 1600 steps/s.

the poles are aligned at each step. The PM-VR hybrid shown in Fig. 20-47*c* has a high skew rate yet retains holding torque when the power is turned off. Motors can be made to rotate in half-steps to increase accuracy. Performance of stepper motors is described by two types of curves: the pull-out torque versus speed curve, as shown in Fig. 20-48; and the holding torque angle curves, as shown in Fig. 20-49. Stepper motors are available with holding torques up to 4000 oz · in.

Hysteresis Motors. By constructing the secondary core of an induction motor of hardened magnet steel, in place of the usual annealed low-loss silicon-steel laminations, the secondary hysteresis loss can be greatly magnified, producing effective synchronous motor action. Such hysteresis motors, having smooth rotor surfaces without secondary teeth or windings, give extremely uniform torque, are practically noiseless, and give substantially the same torque from standstill all the way up to synchronous speed. A hysteresis motor is a true synchronous motor, with its load torque produced by an angular shift between the axis of rotating primary mmf and the axis of secondary magnetization. When the load torque exceeds the maximum hysteresis torque, the secondary magnetization axis slips on the rotor, giving the same effect as a friction brake set for a fixed torque.

Despite the interesting characteristics of this type of motor, it is limited to small sizes, because of the inherently small torque derivable from hysteresis losses. Only moderate flux densities are practicable, owing to the excessive excitation losses required to produce high densities in hard magnet steel, and, therefore, about 20 W/lb of rotor magnet steel represents the maximum useful synchronous power on 60 Hz. Hysteresis motors have found an important use for phonograph-motor drives, their synchronous speed enabling a governor to be dispensed with and freedom from tone waver to be secured.

The Telechron motor, which is so widely used for operating electric clocks, also operates on the hysteresis-motor principle. In the Telechron motor, a 2-pole rotating field is produced in a cylindrical air space, and into this space is introduced a sealed thin-metal cylinder containing a shaft carrying one or more hardened magnet-steel disks, driving a gear train. The 60-Hz magnetic field causes the steel disks to revolve at 3600 r/min, driving through the gears a low-speed shaft, usually 1 r/min, which merges from the sealed cylinder through a closely fitting bushing designed to minimize oil leakage. Although the magnetic field has to cross a very considerable air-gap length and pass through the tin walls of the metal cylinder, the power required to drive a well-designed clock is so small that sample output is obtained with only about 2-W input for ordinary household-clock sizes.

The hysteresis motor has been displaced for phonograph and tape-reel drives by the transistor-driven brushless dc motor. It has been displaced for electric clocks by solid-state circuits with digital readout.

20.6 ALTERNATING-CURRENT COMMUTATOR MOTORS

Classification. As compared with the induction motor, the ac commutator motor possesses two of the advantages of the dc motor: a wide speed range without sacrifice of efficiency and superior starting ability. In the induction motor, the starting torque is limited by the small space-phase displacement between the air-gap flux and the induced secondary current and by magnetic saturation of the flux paths. In the ac commutator motor, on the other hand, the air-gap flux and current are held at the optimum space-phase displacement by proper location of the brush axis, and the secondary current is not limited by magnetic saturation, giving high torque per ampere at starting. Furthermore, the series commutator motor may be operated far above the induction-motor synchronous speed, giving high power output per unit of weight.

Alternating-current commutator motors may be grouped into two classes:

1. Those motors in which the resultant mmf providing the flux increases with the load. When operated from a source of constant voltage, the speed of such motors decreases with increasing load. They are termed *series motors* from the similarity of their characteristics to those of series-wound dc motors. The speed at any given load may be varied by changing the applied voltage or, in some cases, by shifting the brushes.

2. Those motors in which the resultant mmf providing the flux is substantially constant irrespective of the load. For operation from a source of constant voltage, the speed of such motors is approximately constant. The speed may, however, be increased or decreased (independently of the load) by increasing or decreasing the voltage at the terminals of the motor, by brush shifting, or by the provision of suitably disposed and connected auxiliary coils. Such motors are termed *shunt motors*.

Alternating-current commutator motors are either single-phase or polyphase. A unique characteristic of all single-phase motors is a double line-frequency pulsation of the torque produced, corresponding to the sinusoidal variation twice each cycle of the single-phase power supplied. This torque pulsation is partly transmitted to the load, causing small speed pulsations and necessitating special coupling and mounting designs to minimize vibration and fatigue stresses.

Polyphase commutator motors have the advantage of better inherent commutating ability, due in part to the need for shifting the rotor current only 60° in time phase at each brush stud for a 6-phase motor or 30° for 12 phases, as compared with 180° shift for a single-phase or dc machine. Single-phase motors are generally limited to sizes below about 10 hp, except for railway applications.

With the advent of solid-state devices, the ac commutator motor is being displaced by the thyristor-rectifier-powered dc motors and inverter-fed induction motors, at less cost and superior performance. The dc motor does not have the difficulties of commutation, the requirement for extra windings, and shifting brush arrangements, and can be built on an unlaminated frame. The induction motor has no commutator and can run at the high speeds of the commutator motor.

Single-Phase Straight Series Motor. An ordinary dc series motor, if constructed with a well-laminated field circuit, will operate (although unsatisfactorily) if connected to a suitable source of single-phase alternating current. Since the armature is in series with the field, the periodic reversals of current in the armature will correspond with simultaneous reversals in the direction of the flux, and consequently the torque will always be in the same direction. But the inductance of the motor will be so great that the current will lag far behind the voltage, and the motor will have a very low power factor. The entire amount of armature flux produced along the brush axis generates a reactive voltage in the armature, which must be overcome by the applied voltage, without performing any useful function whatever.

When the motor is first thrown in the circuit, and before the armature has moved from rest, the field constitutes the primary of a transformer and sends flux through the armature core. Those armature turns, which at that instant are short-circuited under the brushes, act as short-circuited secondary coils and are traversed by heavy currents, which serve no useful purpose whatever and occasion serious heating. When the armature starts to revolve, these short-circuited turns are opened as they pass out from under the brushes and are replaced by other turns, which are momentarily short-circuited and then opened. These interruptions of heavy currents are accompanied by serious sparking, since the heating is concentrated at the few segments on which the brushes rest. As soon, however, as a certain speed is acquired, the heating is distributed over all the segments and the conditions are ameliorated. This source of sparking is, then, most serious at the moment of starting. This difficulty has been minimized by operating at a lower frequency than 60 Hz and by the employment of leads of high resistance connecting the winding to the commutator segments.

The simple single-phase series motor has therefore two major faults, low power factor and poor commutation at low speeds, confining its use to fractional horsepower and very high speed applications.

Single-Phase Compensated Series Motor. In all except the smallest sizes, it is usual to employ a compensating winding on the stator, in series with the armature and so arranged that its mmf as nearly as possible counteracts the armature mmf. A commutating winding is also frequently used, which somewhat overcompensates the armature reaction along the interpolar, or commutating-zone, axis and so provides a voltage to aid the current reversal, just as in a dc motor. By these means, the flux along the brush axis is reduced to a small fraction of its uncompensated value, and the power factor of the motor is greatly improved. Further improvement of the power factor is secured by using a smaller air gap and correspondingly fewer field ampere-turns than in an uncompensated motor, thus reducing the reactive voltage in the series field to a minimum.

Universal Motors. Small series motors up to about $^1/_2$-hp rating are commonly designed to operate on either direct current or alternating current and so are called *universal motors*. Universal motors may be either compensated or uncompensated; the latter type is used for the higher speeds and smaller ratings only. Owing to the reactance voltage drop, which is present on alternating current but absent on direct current, the motor speed is somewhat lower for the same load ac operation, especially at high loads. On alternating current, however, the increased saturation of the field magnetic circuit at the crest of the sine wave of current may materially reduce the flux below the dc value, and this tends to raise the ac speed. It is possible, therefore, to design small universal motors to have approximately the same speed-torque performance over the operating range, for all frequencies from 0 to 60 Hz. On a typical compensated-type $^1/_4$-hp motor, rated at 3400 r/min, the 60-Hz speed may be within 2% of the dc speed at full-load torque but 15% or more lower at twice normal torque, while on an uncompensated motor the speed drop will be materially greater.

The commutation on alternating current is much poorer than on direct current, owing to the current induced in the short-circuited armature coils, and this provides a definite limitation on their size and usefulness. If wide brushes are used, the short-circuit currents are excessive and the motor-starting torque is reduced, while if narrow brushes are used, there may be excessive brush chatter at high speeds, causing short brush life. Good design, therefore, requires careful proportioning of commutator and brush rigging to meet conflicting electrical, mechanical, and thermal requirements. Universal motors are generally used for vacuum cleaners, portable tools, food mixers, and similar small devices operating at maximum speeds of 3000 to 10,000 r/min.

The speed of the universal motor is controlled by means of a half-wave thyristor, or full-wave triac, as shown in Fig. 20-50. The control device governs the half-wave average voltage applied

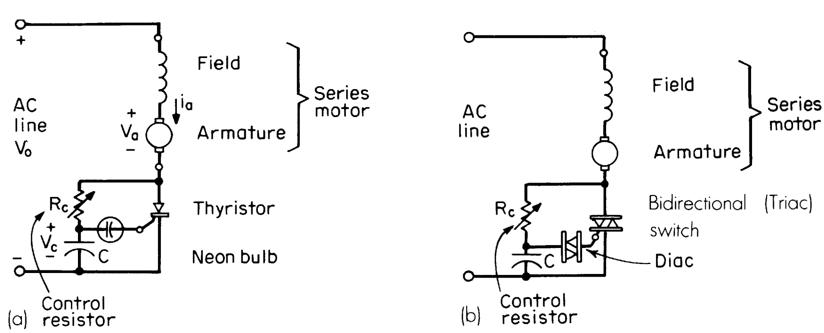


FIGURE 20-50 (*a*) Half-wave series universal motor circuit; (*b*) full-wave series universal motor circuit and the holding torque angle curves, as shown in Fig. 20-49. Stepper motors are available with holding torque up to 400 oz · in.

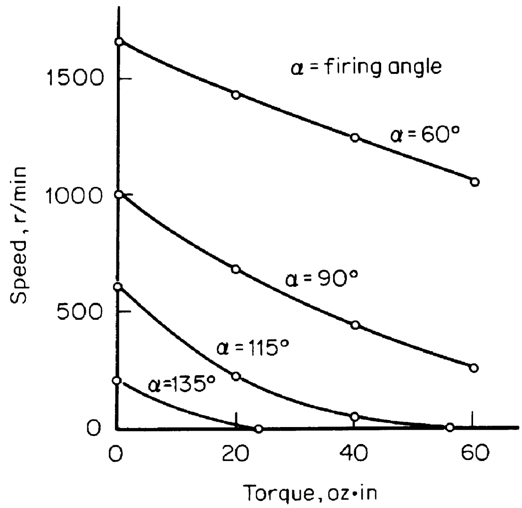


FIGURE 20-51 Measured speed-load torque characteristics of series motor and half-wave thyristor control.

to the motor as a function of the firing angle. The firing circuits are usually relatively simple. The speed is controlled by changing a resistance value, such as R_c, in Fig. 20-50. The characteristics of a universal (series) motor with half-wave control are shown in Fig. 20-51.

20.7 FRACTIONAL-HORSEPOWER-MOTOR APPLICATIONS

Scope. A fractional-horsepower motor is defined by NEMA as either (1) a motor built in a frame with a NEMA two-digit frame number or (2) a motor built in a frame smaller than the NEMA frame for a 1-hp, open-construction, 1700- to 1800-r/min, induction motor. The two-digit frame number is defined as 16D, where D is the height of the shaft centerline above the bottom of the mounting base. Fractional-horsepower motors include 1-phase and 3-phase induction and synchronous motors, 1-phase universal motors, and dc motors. Ratings, with minor exceptions, are $^1/_{20}$ to 1 hp, inclusive. Motors of smaller ratings are classified as subfractional or miniature.

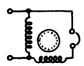
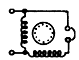
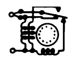
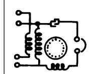
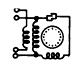
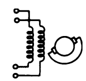
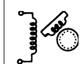
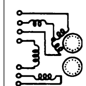
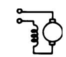
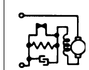
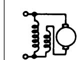
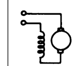
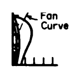
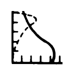
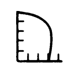
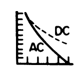
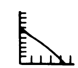
Purpose. General-purpose motors are of open construction, rated at 60°C temperature rise by resistance over a 40°C ambient temperature. They are designed according to standard ratings with standard operating characteristics and mechanical construction for rise under usual service conditions without restriction as to a particular application or type of application. A definite-purpose motor is any motor designed according to standard ratings with standard operating characteristics or mechanical construction for use under service conditions other than the usual or for on a particular type of application.

Selection of Type. The principal characteristics of fractional-horsepower motors are shown in Table 20-11. For any application, the motor selected should meet the application and power supply requirements at the least cost. Numerous trade-offs are possible, for example, speed changing infinite steps compared to continuously adjustable-speed solid-state drives.

Ratings. Standard voltage and frequency ratings are listed in Table 20-12*a* and 20-12*b*.

Service Conditions. General-purpose motors are designed to operate under the usual service conditions of 0 to 40°C ambient temperature, altitude to 3300 ft (1000 M), and installation on rigid mounting

TABLE 20-11 Characteristics of Fractional-Horsepower Motors

	Alternating current									Direct current (cont.)			
	Single-phase motors							1, 2, or 3 phase	Polyphase	DC or ac (60 Hz or less), universal types		Direct current	
	Split-phase types												
	General-purpose	High-torque	Two-speed, pole changing	Capacitor-start	Capacitor (1-value, or perm, split)	Repulsion-start	Shaded-pole	Nonexcited synchronous (reluctance)	Squirrel-cage induction	Without governor	With governor	Shunt or compound	Series
Schematic diagram of connections Arrangements shown are typical or representative; most of the types illustrated have numerous other arrangements which are also used.								Stator winding may be: split-phase, capacitor-start, capacitor, polyphase					
Characteristic speed-torque curves Ordinates are speed: 1 division = for all ac motors, 20% of syn. r/min; for universal motors, 1000 r/min; for dc motors, 20% of full-load rpm. Each abscissa division = 100% of full-load torque.													
Rotor construction	Squirrel-cage	Squirrel-cage	Squirrel-cage	Squirrel-cage	Squirrel-cage	Drum-wound; commutator	Squirrel-cage	Cage, with cutouts	Squirrel-cage	Drum- wound; commutator	Drum-wound; commutator	Drum-wound; commutator	Drum-wound; commutator
Built-in automatic starting mechanism	Centrifugal switch	Centrifugal switch	Centrifugal switch	Centrifugal switch	None required	Short-circuit	None	Depends on stator winding	None	None	None	None	None
Horsepower ratings commonly available	$^1/_{20}$–$^1/_3$	$^1/_6$–$^1/_3$	$^1/_8$–$^3/_4$	$^1/_8$–$^3/_4$	$^1/_{20}$–$^3/_4$	$^1/_8$–$^3/_4$	$^1/_{2000}$–$^1/_6$	$^1/_{3000}$–$^1/_3$	$^1/_6$–$^3/_4$	$^1/_{150}$–1	$^1/_{50}$–$^1/_{20}$	$^1/_{30}$–$^3/_4$	$^1/_{25}$–$^1/_{30}$
Usual rated full-load speeds (for 60-Hz ac motors; also dc motors)	3450, 1725, 1140, 865	1725	1725/1140 1725/865	3450, 1725, 1140, 865	1620, 1080, 820	3450, 1725, 1140, 865	1450, –3000	3600, 1800, 1200, 900	3450, 1725, 1140, 805	3000–11,000	2000–4000	3450, 1725, 1140, 865	900–2000
Speed classification	Constant	Constant	Two-speed	Constant	Constant, or adjustable varying	Constant	Constant, or adjustable varying	Absolutely Constant	Constant	Varying, or adjustable varying	Adjustable	Constant, or adjustable varying	Varying, or adjustable varying
Means used for speed control			Two-speed switch		Two-speed switch or auto-transformer		Choke or resistor			Choke or resistor	Adjustable governor	Armature resistance	Resistor

Comparative torques: Locked-rotor	Moderate	High	Moderate	Very high	Low	Very high	Low	Low	Very high	Very high	Very high	Very high	Very high
Comparative torques: Breakdown	Moderate	High	Moderate	High	Moderate	High	Low	Moderate	Very high				
Radio interference, running	None	None	None	None	None	None	None	None	None	Continuous	Continuous	Continuous	Continuous
During acceleration	One click	One click	Two click	One click	None	Continuous	None		None	Continuous	Continuous	Continuous	Continuous
Approximate comparative costs between type, for same horsepower rating: Below $^1/_{20}$ hp													
$^1/_2$–$^1/_4$ hp	100	75	210	125	140	128	100	200–200	165–195	75	75	175–225	185
$^1/_{20}$–$^1/_4$ hp	80	54	150	100	100–110	100		275	100	105–175	140–160	120–140	
General remarks: Standard motors are ordinarily designed to operate in ambient temperatures from 10 to 40°C. Variations in line voltage of plus or minus 10%, or variations in frequency of plus or minus 5% are allowable. Locked-rotor currents for single-phase motors, except split-phase high-torque and synchronous types, usually do not exceed the following limits established by NEMA:	For constant-speed operation, even under varying load conditions, where moderate torques are desirable or mandatory, this type is often used in preference to the more costly capacitor-start motor. Meets NEMA starting currents. Typical applications: blowers; centrifugal pumps; duplicating machines; refrigerators; oil burners; unit heaters.	High locked-rotor currents (in excess of NEMA) limit the use of this type on lighting circuits to applications where the motor starts only very infrequently, because of a tendency to cause flickering of the lights. Principal applications: washing and ironing machines; cellar-drainer pumps; tools for a home workshop.	Used where two definite speeds independent of load are required. Ratings above 1/4 hp usually made capacitor-start. Motor shown always starts on high-speed connection; transfer to low speed made by starting switch. Common applications: belted blowers for warm-air furnaces or for other purposes; attic ventilators; air conditioning apparatus.	A general-purpose motor suitable for most applications requiring constant speed under varying loads, high starting and running torques, high overload capacity. Also available as two-speed pole-changing motor above 1/4 hp. A few important applications are: refrigeration and air conditioning compressors; air compressors; stokers; gasoline pumps.	Primarily used for unit heaters, or for other shaft-mounted fans. Essentially a constant-speed motor, but by means of a two-speed switch, or by means of an auto-transformer, other speeds can be obtained, *with fan loads,* of horsepower rating selected closely matches the fan load. Can also be made in intermittent ratings for plug-reversing service.	A constant-speed motor suited to general-purpose applications requiring high starting torque, such as pumps and compressors. An associated type, the repulsion induction (buried cage) is used for door openers and other plug-reversing applications. Has been displaced for many applications by the capacitor-start motor.	For ratings below 1/20 hp, this is a general-purpose motor. For fan applications, speed control is affected by use of a series choke or resistor. Applications: fans, unit heaters, humidifiers, hair driers, damper controllers.	Cutouts in rotor result in synchronous-speed characteristics. Curve shown is for split-phase stator. Pull-in ability is affected by inertia of connected load. Used for teleprinters, facsimile-picture transmitters, graphic instruments, etc. Clocks and timing devices usually use shaded-pole hysteresis motors rated at a few millionths of a horsepower.	Companion motor to capacitor-start motor with comparable torques and generally suited to same applications if polyphase power is available. Inherently plug-reversible and suitable for door openers, hoists, etc. High-frequency motors used for high-speed applications, as for wood-working machinery, rayon spinning, and portable tools.	Light weight for a given output, high speeds, varying-speed and universal characteristics make this type very popular for hand tools of all kinds, vacuum cleaners, etc. Ratings above $^1/_4$ hp usually compensated. Some speed control can be effected by a resistor or by use of a tapped field. Used with reduction gear for slower speed applications.	By means of a centrifugal governor, a constant-speed motor having the advantages of the universal motor is obtained. Governor may be single-speed or adjustable even while running. Speed is independent of applied voltage. Used in typewriters, calculating machines, food mixers, motion-picture cameras and projectors, etc.	A constant-speed companion motor for the capacitor-start or split-phase motor for use where only dc power is available. For unit-heater service, armature resistance is used to obtain speed control. Not usually designed for field control.	Principally used as the dc companion motor to the shaded-pole motor for fan applications. Used in these small ratings in place of shunt motors to avoid using extremely small wire.

	Amperes at	
Rating, hp	115 V	230V
$^1/_8$ and smaller	20	10
$^1/_4$	23	$11^1/_2$
$^1/_3$	31	$15^1/_2$
$^1/_2$	45	$22^1/_2$
$^3/_4$	61	$30^1/_2$

Fractional horsepower motors are built for across-the-line starting.

The standard direction of rotation is counterclockwise facing the end opposite the shaft extension.

TABLE 20-12 Voltage Ratings of (a) AC Fractional-Horsepower Motors and (b) DC Fractional-Horsepower Motors.

Motors	Frequency, Hz	Voltage
Single-phase	60	115, 230
	50	110, 220
3-Phase	60	115, 200
		230, (460)
	50	220, 380
Universal	60*	115, 230

*Can operate from dc to 60 Hz.

(a)

Primary power source	Rating, hp	Armature voltage	Field voltage
Los-ripple dc	$^1/_{20}-1$	115, 230	115, 230
1-Phase rectifier	$^1/_{20}-^1/_2$	75	50, 100
		90	50, 100
		150	100
	$^3/_4-1$	90	50, 100
		180	100, 200
3-Phase rectifier	$^1/_4-1$	240	100, 150
			240

(b)

surfaces where there is no interference with the ventilation. Some general-purpose, definite-purpose, and special-purpose motors can operate under one or more unusual service conditions, which include exposure to dust, lint, fumes, radiation, steam, fungus, shock; operation where voltage, frequency, waveform, and form factor deviate from standards; and overspeed, overtemperature, and excess altitude operation. The manufacturer should be consulted for operation under unusual service conditions.

Thermal Protection. Many single-phase motors are now available with a built-in thermal protector, which affords complete protection from burnout due to any type of overload, even a stalled rotor. Most such devices are automatic-resetting, but some are manual-resetting. Motors that are protected usually are marked externally in some way to indicate the fact.

Reversibility. In general, standard motors of the types listed in the table can be arranged by the user to start from rest in either direction of rotation. There are exceptions, however. Shaded-pole motors, unless of a special design, can be operated in only one direction of rotation. Small dc and universal motors often have the brushes set off neutral, preventing satisfactory operation in the reverse direction. Single-phase motors, which use a starting switch ordinarily cannot be reversed while running at normal operating speeds, because the starting winding, which determines the direction of rotation, is then open-circuited. By use of special relays this limitation of split-phase and capacitor-start motors can be overcome when necessary. Such motors are built for small hoists. High-torque intermittent-duty permanent-split capacitor motors; repulsion-induction (buried-cage) motors; and split-series dc or universal motors are often built for plug-reversing service. Standard polyphase induction motors can be reversed while running, as can the smaller ratings of dc motors; such applications should preferably be taken up with the motor manufacturer.

Mechanical Features. Rigid and rubber-mounted motors are commonly available. Sleeve and ball bearings are both standard. Sleeve-bearing motors are designed for operation with the shaft horizontal, but ball-bearing motors can be operated with the shaft in any position. For operation with the shaft vertical, sleeve-bearing motors may require a special design. Rubber mounting is widely used for quiet operation, because all single-phase motors have an inherent double-frequency torque

TABLE 20-13 Approximate Starting and Full-Load Current for Single-Phase 115-V Motors

Rating,	Max. locked-rotor current, A		3450 r/min		1725 r/min		1140 r/min		865 r/min	
hp	Des. O	Des. N	A	W	A	W	A	W	A	W
$^1/_8$	50	20	2.9	207	2.7	176	3.9	207	5.4	245
$^1/_6$	50	20	3.2	254	3.0	214	4.3	254	6.0	296
$^1/_4$	50	26	4.2	352	3.9	301	5.6	352	8.1	414
$^1/_3$	50	31	5.3	460	4.9	395	7.0	460	9.8	540
$^1/_2$	50	45	7.4	678	6.9	574	9.8	678		
$^3/_4$	. . .	61	10.6	981	9.9	835				
1	. . .	80	13.3	1260						

Note: A = amperes; W = watts; Des. = Designation.

pulsation. An effective and common arrangement uses rubber rings concentric with the shaft and so arranged as to provide appreciable freedom of torsional movement but little other freedom. Sometimes the driven member picks up the double-frequency torque pulsation and amplifies it to an objectable noise, for example, a fan with large blades mounted rigidly on the shaft. The cure for this difficulty is an elastic coupling between the shaft and the driven member; no amount of elastic suspension of the stator can help. Standard motors are generally open and of drip-proof construction. Splashproof and totally enclosed motors are easily available.

Inputs of Small Single-Phase Motors. See Table 20-13. Full-load torque, in terms of horsepower and rated speed, is

$$\text{Full-load torque, oz}\cdot\text{ft} = \frac{84{,}000 \times \text{hp}}{\text{r/min}} \tag{20-33}$$

Application Tests. The primary object of any application test is to determine the power requirements of the appliance or device under various significant operating conditions. A convenient way of doing this is to use a motor of approximately the right horsepower rating and of predetermined efficiency at various outputs. Watts input are carefully measured under each condition. From the watts input observed (never use current as a measure of load except for dc motors) and the known efficiency, the load is readily determined. Care should be taken in measuring the watts input to correct for the meter losses.

A second, and equally important, object of the test is to determine the actual locked-rotor and pull-up torques required by the appliance. The locked-rotor and pull-up torques of the test motor should be known or measured at rated voltage and frequency. (Locked-rotor torque often varies with slight changes in rotor position.) Using a transformer or induction regulator to obtain a variable voltage (do not use a resistance or choke for this purpose), measure the minimum voltage at which the motor will start the appliance and also the minimum voltage at which it will pull it up through switch-operating speed. Assuming that the pull-up and locked-rotor torques each vary as the square of the applied voltage, it is then a simple matter to determine the actual locked-rotor and pull-up torques required by the device. After a motor has been selected, it should be determined whether it can operate the device at 10% above and below normal rated voltage of the motor or over a wider range of voltage, if desired. If exceptional load conditions may occasionally be encountered, use of a motor equipped with inherent-overheating protection is often desirable.

Definite-Purpose Motors. For a number of important applications, involving large quantities of motors, NEMA has developed standards to meet these special requirements effectively and economically. Motors built to these standards are usually more readily obtainable and economical than special motors tailored to one application. Highlights and distinguishing features are given in Table 20-14. More details can be obtained in NEMA Standards.

TABLE 20-14 NEMA Standards for Definite-Purpose Motors

Application	Principal types	Distinguishing features
Universal motors	Universal: salient-pole and distributed field	Dimensional standards; common practices utilizing parts
Hermetic motors	Split-phase, capacitor-start, polyphase	Parts only for hermetic refrigeration condensing units
Belt-drive refrigeration compressors	Capacitor-start, repulsion-start, polyphase	Open; sleeve bearings, extended rear oiler; automatic-reset thermal overload protection
Jet-pump motors	Split-phase, capacitor-start, repulsion-start, polyphase	3450 r/min; ball bearings; open; machined back-end shield; automatic-reset overload protection
Motors for shaft-mounted fans and blowers	Split-phase, permanent-split capacitor, polyphase	Enclosed; horizontal; sleeve bearings; vertical, ball bearings; extended through bolts; capacitors on front end shield
Shaded-pole motors for shaft-mounted fans and blowers	Shaded-pole; two-speed, three speed	Open or totally enclosed; sleeve bearings; high slips
Belted fans and blowers	Split-phase, capacitor-start, repulsion-start; two-speed split-phase and capacitor-start	Open; sleeve bearings; resilient mounting; automatic-reset overload protection; extended rear oiler
Stoker motors	Capacitor-start; repulsion-start; polyphase	Totally enclosed recommended; automatic-reset overload protection
Motors for cellar drainers and sump pumps	Split-phase	Vertical, dripproof, 50°C; two ball bearings, or one ball, one sleeve; mounts on support pipe; built-in float-operated line switch; overload protection
Gasoline-dispensing pumps	Capacitor-start, repulsion-start, polyphase	Explosionproof; sleeve bearing; built-in line switch and capacitor; voltage-selector switch on single-phase
Oil-burner motors	Split-phase	Enclosed, face-mounted, round-frame; manual-reset overload protection; two line leads
Motors for home-laundry equipment	Split-phase	Low-cost, high starting current; open, 50°C; round-frame with ungrounded mounting rings; shaft extension with flat and hole for coupling
Motors for coolant pumps	Split-phase, capacitor-start, repulsion-start, polyphase	3450- and 1725-r/min; totally enclosed; ball bearings; machined back-end shield
Submersible motors for deep-well pumps	Split-phase, capacitor, polyphase	3450 r/min; designed for operation totally submerged in water not over 25°C (77°F); use external relay for starting

Small Synchronous Motors. Small synchronous motors in the 1.5- to 25-W range for timing, tape drives, small fans, and record players are available as brushless dc motors or as hysteresis motors. The brushless dc motor consists of a permanent-magnet field, 2-phase, synchronous motor driven by transistors from a dc source. The transistors are switched from a Hall-device signal that senses the rotor position. A regulator maintains constant speed.

Shaded-pole hysteresis motors, which operate at synchronous speed, are essentially the same as shaded-pole induction motors except that they use rotors of hardened-steel rings of a material having high hysteresis loss. Large quantities of such motors are built for clocks and timing devices. Clock motors have an input of 1.5 to 2 W and an output of a few millionths of a horsepower. Large motors with inputs up to 15 W are built for heavier duty applications. Rotor speeds are commonly 450, 600, and 3600 r/min. Most of these motors are furnished with built-in reduction gears to give output speeds of 60 r/min to 1 r/month.

Reluctance motors, both self-starting and manual-starting types, are available for similar applications. Another type used is the synchronous-inductor motor, which is essentially an inductor alternator used as a motor; field excitation is furnished by a permanent magnet.

20.8 MOTOR CONTROL

Industrial motor control includes (1) motor-starting devices, (2) speed-control devices, (3) stopping devices, and (4) motor-protecting devices.

Industrial motor control is designed and built in accordance with rules and standards established by several organizations. Detailed design-construction and test information is contained in such publications as the *National Electrical Code,* National Board of Fire Underwriters; *Standards for Industrial Control Apparatus,* Institute of Electrical and Electronics Engineers; *Industrial Control Standards,* National Electrical Manufacturers Association; *Standard for Industrial Control Equipment,* Underwriters' Laboratories; and *Standard Rotation, Connections and Terminal Markings for Electric Power Apparatus,* American National Standards Institute.

The essential functions of motor control are the starting, speed regulating, stopping, and protecting of electric motors.

20.9 MOTOR-STARTING DEVICES

A *contactor* is a device, generally magnetically actuated, for repeatedly establishing and interrupting an electric power circuit.

Figure 20-52 illustrates a single-pole dc contactor. Contactors of this type are rated on a continuous-current-carrying-capacity basis and on an intermittent-duty basis, at values depending on the duty cycle. The NEMA Standard 8-h open ratings range from 25 to 2500 A. The intermittent ratings are $133^1/_3\%$ of the open ratings. The shunt-operating coil is designed to withstand 110% of rated voltage continuously and to close the contactor successfully at 80% of rated voltage.

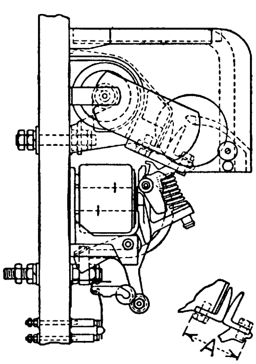


FIGURE 20-52 Direct-current contactor.

The *magnetic blowout* consists of a coil wound on a steel core and mounted between steel pole pieces. The pole pieces are lined with refractory material. The assembly is enclosed in an insulated box, which is swung down over the contacts. The blowout coil is generally connected in series with the contactor and carries motor current with the contactor closed. The current sets up a magnetic field through the core and pole pieces of the blowout structure and across the contact tips. When an arc is formed, the magnetic field of the arc and the magnetic field of the blowout repel each other and the arc is forced upward and away from the contacts. The extinguishing action, due to the lengthening of the arc and the cooling of the refractory material, is extremely rapid and thereby greatly reduces the wear and burning of the contacts.

Several factors are important in the performance of contactors. To obtain trouble-free service and maximum contact life, the following items should be in accordance with the manufacturer's specifications: initial and final contact pressures, magnetic gap, arc gap, and wear allowance. The contact pressures can be measured by means of a spring balance, initial pressure with the contactor open,

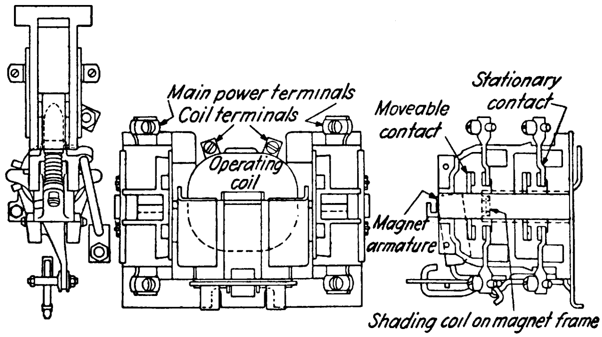


FIGURE 20-53 Alternating-current contactor.

and final pressure with the contactor closed. The magnetic gap is the distance from the centerline of the core to a corresponding point on the armature lever, and the arc gap is the distance between the arcing tips. Contacts can be kept smooth by filing with a fine file. Contact-wear allowance is the total thickness of material that may be worn away before the contact between the two surfaces becomes ineffective. The contacts should be renewed when worn so that the distance A (Fig. 20-52) between the back edges of the contacts, with the contactor closed, becomes less than the specified amount. This usually corresponds to the condition where the contacts are worn to approximately half their original thickness.

Current-carrying contacts are usually made from copper, either left plain or plated with silver or cadmium. The contacts should not be lubricated. Hinge pins and bearings should be lubricated with a light machine oil. The surface of the core and the armature, which seal when the contactor is closed, should be kept clean.

An ac contactor is similar in construction to a dc contactor, except that laminated iron structures are used. A shading coil is used at the core face to ensure continuous force of the armature and thus obtain quiet operation. AC contactors are available with 2, 3, or 4 main poles for interrupting all line circuits to single-phase, 3-phase, or 2-phase 4-wire motors (see Fig. 20-53 and Table 20-15). Standard ac contactors are designed to interrupt 10 times rated motor current, based on the contactor horsepower rating. The contactors are also designed to thermally withstand 15 times rated motor

TABLE 20-15 Typical Ratings of AC Contactors Used as Across-the-Line Magnetic Starters with 3-Phase Motors

Contactor size	Rating, A	Horsepower at 110 V	Horsepower at 220 V	Horsepower at 440–550 V
00	9	$^3/_4$	$1^1/_2$	2
0	18	2	3	5
1	27	3	$7^1/_2$	10
2	45		15	25
3	90		30	50
4	135		50	100
5	270		100	200
6	540		200	400
7	810		300	600
8	1215		450	900
9	2250		800	1600

current for 1 s, to permit protective devices such as circuit breakers and fuses to clear fault current carried by the contactor.

A *drum switch* consists of stationary contact fingers held by spring pressure against contact segments on the periphery of a rotating cylinder or sector. Drum controllers have many advantages over the faceplate and multiple-switch types. The mechanical construction is better, heavy contact pressures can be maintained, parts can be well insulated, blowout magnets and arc shields can be used, and the structure can easily be completely enclosed. Less space is required by the drum control, and it is easier to operate. Drum controllers are built in 8-h ratings for dc motors up to 40 hp, 115 V, and 75 hp, 230 V.

The design of starting-duty resistors requires the determination of the total ohms, the distribution of this resistance between the steps available, and the calculation of the current-carrying capacity, and the selection of the resistance material. Standard resistors to meet various classes of service are designated by class numbers in accordance with the NEMA *Table of Classification of Resistors*. NEMA also publishes a resistor application table intended as a guide in specifying and designing resistors (see *NEMA Industrial Control Standards*). This table lists typical machines with the corresponding NEMA resistor-classification number. For example, a lathe should have a dc starter with resistor classification No. 115. This resistor has sufficient total ohms to limit the current inrush on starting to 150% of full-load current. It will be designed to have current-carrying capacity for an average accelerating current (rms value) of 125% full-load current, on the basis of starting once during each 80-s period and with an accelerating time of 5 s.

Resistors must be available in a wide range of ohmic values and current capacity. Resistor units are stacked in parallel and series combinations to achieve the required values. For low ohms and high capacity, cast-iron or punched-steel grids are used. These steel grids range from 0.01-Ω, 160-A to 0.40-Ω, 20-A continuous rating per grid. For high ohms and low capacity, wire-wound resistors are used with a unit resistance range from 4.0 to 6400 Ω and dissipation up to 900 W. Intermediate resistance requirements are fulfilled with edge-wound ribbon resistors with a unit resistance range of 0.05 to 8.6 Ω and dissipation up to 1320 W.

Resistors must be sized and arranged in assemblies so that the temperature rise for bare resistive elements does not exceed 375°C above 40°C ambient. The maximum current should not exceed the 10-s rating. Care must be taken to limit excess voltage on wire-wound resistors to avoid surface flashover. Resistors are generally insulated for 600 V rms to ground. Manufacturers will provide assembled resistor units for specific functions or provide the derating curves for the arrangement of the resistor units into assemblies, as a function of duty cycle, ambient temperature, and grouping.

20.9.1 AC Motor Starting

Selection of an ac starter is a compromise between requirements and cost. The primary requirements of the starter, obviously, are that the motor starting torque shall be adequate to start the load under worst-case line voltage and load conditions; also, that the line current shall not exceed limits set by the utility or plant voltage dip. A useful table is shown in Table 20-16. The available starters are listed in descending value of starting torque, based on a 60-hp, 440-V, 60-Hz, 900-r/min motor. Compared to the starters with series-connected elements, the autotransformer starter provides a means for reducing the line current below the motor current. The current and torque in Table 20-16 are shown for 100% line voltage. For reduced line voltage, the current is reduced in proportion; the torque is reduced as the square of the voltage. Limit on voltage dip is 15% to 20%.

The secondary requirements in starter selection include smoothness of acceleration, maintenance, power factor, reliability, and efficiency. The selection of a closed-transition starter depends upon whether the motor and the supply line can withstand the peak current as the time the starter transfers the motor to full voltage.

Alternating-current across-the-line starters are simple in construction, easy to install and maintain, and inexpensive. A typical starter consists of a 3-pole contactor with a thermal overload relay for protecting the motor. The starter connects the motor directly to the line, impressing full voltage to the motor terminals. It is particularly suitable for squirrel-cage motors. Since these starters connect the motor directly to the supply lines, the motor will draw an inrush current of 6 to 10 times running current. In the

TABLE 20-16 Comparison of Methods for Starting 3-Phase Motors (Example of 60-hp, 900-r/min, 60-Hz motor)

Method of starting	Starting current drawn from the line as a percentage of full-load current	Starting torque as a percentage of full-load torque
Connecting motor directly to the line full potential	470	160
Autotransformer 80% tap	335	105
Resistor starter to give 80% applied voltage	375	105
Part winding	235	70
Autotransformer 65% tap	225	67
Resistor starter to give 65% applied voltage	305	67
Solid-state starter	300	65
Star-delta starter	158	54
Resistance starter to give 58% applied voltage	273	54
Autotransformer 50% tap	140	43
Resistor starter to give 50% applied voltage	233	43

majority of installations this is not objectionable and will not damage the motor or the driven machinery. When the starting inrush must be lower, some form of reduced-voltage starting must be used. The common types of starters are autotransformer, primary-resistance, part winding, Y-Δ, and solid-state.

Autotransformer starters have two autotransformers connected in open Δ to provide reduced-voltage starting. Three taps are supplied, as shown in Fig. 20-54, giving 50%, 65%, and 80% of full line voltage. The motor current varies directly as the voltage impressed on the motor terminals. The line current varies as the square of the impressed voltage and is therefore lower than with resistor-type starters. The torque also varies as the square of the impressed voltage. The 50% voltage tap will therefore provide 25% starting torque. Connections should be made to the lowest tap that will give the required starting torque.

Characteristics of this type of starter are low line current, low power from the line, and a low power factor. Acceleration is not continuous, because the torque developed by the motor remains practically constant during the starting period, on the first step; then changes to another value on the second step. The starter shown in Fig. 20-54 is an open-transition type. The motor is disconnected from the line during the transfer period. A closed-transition starter which uses the Korndorfer connection is shown in Fig. 20-55. When the *S* contactor opens, the autotransformers act as series reactors until the *R* contactor closes.

Autotransformer starters are available in manual and automatic types. In the manual type the contacts are operated by means of a lever extending to one side of the enclosing case. The lever is equipped with a low-voltage release magnet. The automatic open-transition starter (Fig. 20-54) consists of a 5-pole starting contactor *S* and a 3-pole running contactor *R*. The closed-transition starter (Fig. 20-55) consists of a 3-pole main contactor (*M*) and 2-pole start and run contactors (*S* and *R*).

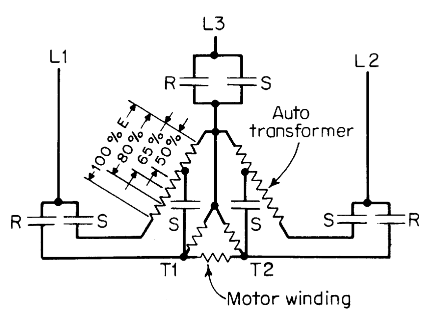


FIGURE 20-54 Connections for autotransformer starter.

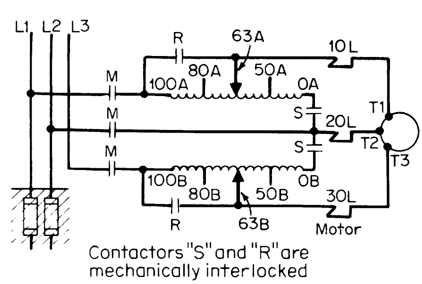


FIGURE 20-55 Connections for a closed-transition autotransformer motor starter.

When the "run" button is pressed, the starting contactor closes, connecting the transformer to the line and the motor to the reduced-voltage taps. A timing relay is operated by the starting contactor. The motor accelerates, and, after a specified number of seconds, the timer contacts close, deenergizing the starting contactor and energizing the running contactor. The transformer is disconnected and the motor connected to line voltage.

Primary-resistor-type starters connect the motor to the line through a series resistor. Reduced voltage at the motor is obtained because of the voltage drop across the resistor. As the motor accelerates, the current drawn from the line declines, and consequently the voltage drop across the resistor is lowered, and the motor voltage at the motor terminals is increased. The torque delivered by the motor is therefore constantly increased as the motor speed increases. After a definite interval, a timing device operated by the main contactor energizes the accelerating contactor, which short-circuits the resistor. There is no transfer period during which the motor may lose speed, and therefore smooth acceleration is obtained. In comparison with the autotransformer-type starter, the primary-resistor type takes more power from the line on starting but provides smoother acceleration, faster acceleration with a given initial torque, and higher power factor. In the smaller sizes, the primary-resistor starter costs less than the autotransformer starter.

In *part-winding ac starters,* the motor winding must be in two parts, and at least six terminal leads must be provided on the motor. The method is therefore applicable to those motors which are designed for use on either of two voltages, the windings being in parallel on the lower voltage and in series on the higher voltage. For example, a 230/460-V motor could be used on 230 V with a part-winding controller. The controller would then be arranged to connect one section of the winding to the supply lines as soon as the starting button is pressed. Then, after a time delay provided by a timing relay, a second contactor would connect the other section of the motor winding to the supply lines, in parallel with the first section. In this way, the starting current is reduced to approximately one-half of what would be required if both winding sections were connected at the same time, as they would be with a standard 3-lead motor. The starting torque when the first winding section is connected will be less than half of the torque that would be obtained if both sections were connected at the same time. Contactors used for part-winding starters need capacity to handle only the circuit which they control, and so may be rated at one-half of the rating that would be required to handle the whole motor. Overload relays are provided for each section of the winding.

Y-Δ *ac starters* are a form of reduced voltage starter used with 6-lead motors in which 57% voltage is applied to the windings on the first step, full voltage on the second step. The starting current and starting torque are 33% of the full-voltage values. The Y-Δ starter is used for compressors and other loads that can be unloaded for starting, or can tolerate the 33% starting torque.

Y-Δ starters are built for open-transition, as shown in Fig. 20-56, or closed-transition operation, as shown in Fig. 20-57. The open-transition starter operates as follows: Relay contactor *S* is energized, connecting the motor windings in Y. A normally open auxiliary contact on contactor *S* closes,

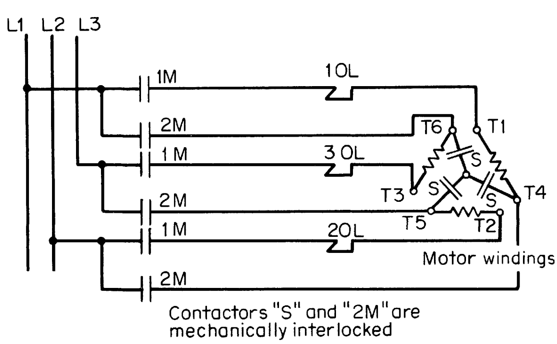


FIGURE 20-56 Open-transition Y-Δ starter.

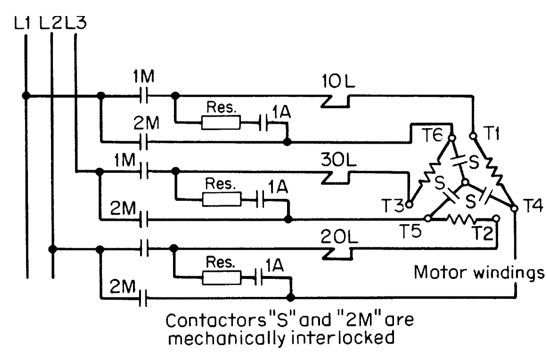


FIGURE 20-57 Closed-transition Y-Δ starter.

energizing contactor 1*M*. Its contacts close, energizing the motor windings in Y. After a predetermined interval, the contactor is deenergized, contacts on timer *TR* open deenergizing contactor *S*, opening its contacts and thereby opening the Y-connected winding. The motor is now temporarily deenergized. A normally closed auxiliary contact on contactor *S* closes, energizing contactor 2*M* in addition to 1*M*. The motor is reenergized in Δ.

The closed-transition starter of Fig. 20-57 operates as follows: Contactor *S* is energized, connecting the motor windings in Y. A normally open auxiliary contact on contactor *S* closes, energizing contact 1*M*, closing its contacts, energizing the motor windings in Y. After a predetermined interval, contactor 1*A* is energized, connecting resistors *Res* in Y and paralleling them across the Y-connected motor winding. A normally closed auxiliary contact on contactor 1*A* opens, deenergizing contactor *S*, opening its contacts and placing resistors *Res* in series with the motor winding. The motor is now connected in Δ. A normally closed auxiliary contact on contactor *S* closes, energizing contactor 2*M*, closing its contacts and thereby shorting out the resistors *Res*. The Δ-connected compressor motor is now energized at full voltage.

Solid-state ac starters employ back-to-back phase-controlled thyristors in two or three of the lines to the motor as shown in Fig. 20-58. The thyristors are controlled during the starting period to maintain about 300% line and motor current by gradually increasing the motor voltage from the initial value. Starting is smooth; the current and starting torque can be adjusted easily. The solid-state starter is applied where the line current is critical and where repetitive motor starting limits the life of electromagnetic contactors.

Slip-ring ac motor starters consist of a contactor to connect the motor primary to the supply lines and a resistor and resistor switching means for the secondary circuit. The starting torque depends on the ohmic value of resistance used; maximum torque is obtained when the resistance is selected for an inrush of approximately 3 times full-load current. Sufficient resistance is generally used to limit the inrush current to 150% or 200%. The resistor is cut out step by step as the motor accelerates, until the slip rings are short-circuited. The commutating means may be a faceplate controller, a drum, or a series of magnetic contactors controlled by current or time relays. High starting torque and low running slip can be obtained with a slip-ring motor.

20.9.2 DC Motor Starting

Direct-current motors of small capacity may be started by connecting the motor directly to line voltage. Motors rated 2 hp or more generally require a reduced-voltage starter. The reduced voltage for starting is obtained by using resistance in series with the motor armature or by varying the armature supply voltage. Manual or magnetic control may be used.

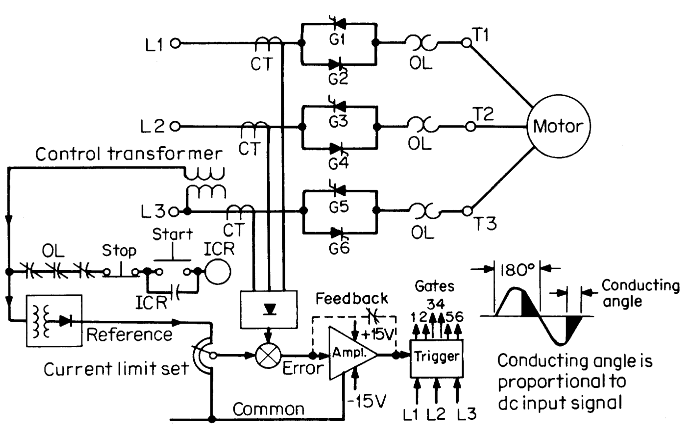


FIGURE 20-58 Diagram of solid-state ac motor starter.

DC motors in adjustable-voltage, adjustable-speed drives are started by turning the speed control up from zero to the desired speed, or by internal circuits that ramp the armature voltage to the desired value. Starting equipment, other than the armature-voltage rectifier or generator, is not required.

Direct-current manual starters are satisfactory for applications that do not require frequent starting and stopping and where the starter can be mounted near the operator without requiring long motor leads. Across-the-line starters provide the simplest means of starting small dc motors. Manually operated switches for this service are available in sizes up to 1.5 hp at 115 V and 2 hp at 230 V. For larger motors resistance is connected in series with the motor armature to limit the current inrush on starting. A manually operated means is then provided for removing the resistor from the circuit in a series of steps. Starters are available in the faceplate type, the multiple-switch type, and the drum type. The faceplate type is built for motors up to 35 hp, 115 V, and 50 hp, 230 V. It consists of a movable lever and a series of stationary contact segments to which sections of resistor are connected. The resistor sections are short-circuited one at a time by moving the lever across the segments.

Manual starters have generally been replaced by push-button-operated magnetic control that incorporates overload protection and other safety features.

Direct-current magnetic starters are used for applications where case and convenience of operation are important; where the starter is operated frequently; where the motor is located at a distance from the operator; where automatic control by means of a pressure switch, limit switch, or similar device is desired; and for large motors which require the switching of heavy currents. Resistance is connected in series with the motor armature to limit the initial current and is then short-circuited in one or more steps.

For larger motors a series of magnetic contactors is used, each of which cuts out a step of armature resistance. The magnetic contactors are operated as the motor starts by one of two methods called *current-line acceleration* and *time-limit acceleration;* the starting time is always matched to the burden of the load. Time-limit acceleration is advantageous where the starting time of the motor must be integrated into a timing sequence for an overall machine or process. Examples of each will be given.

Figure 20-59 shows a type of *time-limit acceleration* where the operation of contactors, and therefore the rate of acceleration, is governed by a *magnetically operated definite time relay*. This time relay operates on the principle of discharging a capacitor, thus obtaining a definite time period, which is unaffected by changes in temperature and load or by dust and dirt. With the motor at rest, a circuit is obtained through a normally closed contact on *M* to energize the *CT* timing-delay coil and to charge capacitor *C*1. Contacts *CT*1 and *CT*2 on relay *CT* are open with the relay energized. Capacitor *C*2 is charged through the normally closed contact on the 2*A* contactor. Pressing the "start" button energizes the main contactor *M*, which maintains itself through a normally open interlock finger. Relay *FA* is energized, and its contact *FA*1 short-circuits the field rheostat. The motor accelerates from rest to a speed determined by the value of the *R*1-*R*3 resistor. The circuit to timing relay *CT* is opened by the interlock on *M*, and capacitor *C*1 discharges through the *CT* coil and the *AB* resistor. Contacts *CT*1 and *CT*2 can be individually adjusted to close at any time during the capacitor discharge period. Closing *CT*1 energizes the 1*A* contactor, which short-circuits the *R*1-*R*3 resistor. The motor then accelerates to a speed determined by the value of the *R*2-*R*3 resistor step. Closing *CT*2 energizes 2*A* and connects the motor across the line, permitting it to accelerate to normal speed.

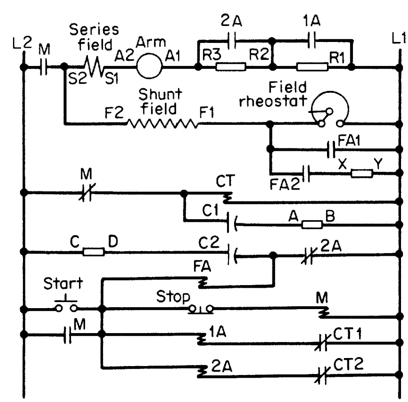


FIGURE 20-59 Time-limit acceleration with definite-time relay.

Relay *FA* is deenergized when 2*A* closes. Contacts *FA*1 and *FA*2 remain closed for a definite time because of the discharge of capacitor *C*2 through the *FA* relay coil. When contact *FA*1 opens, a resistance is inserted in the motor shunt field, equivalent to the field-rheostat resistance and *XY* resistance in parallel. Opening *FA*2 disconnects *XY*, and the motor runs at a speed determined by the setting of the field rheostat.

Direct-current starters with current-limit acceleration are designed to half the starting operation whenever the required starting current exceeds an adjustable predetermined value, the starting operation being resumed when the current falls below this limit. With current-limit acceleration, the time required to accelerate will depend entirely upon the load. When the load is light, the motor will accelerate rapidly, and when it is heavy, the motor will require a longer time to accelerate. For this reason a current-limit starter is not so satisfactory as a time-limit starter for drives having varying loads. Time-limit starters are simpler in construction, accelerate a motor with lower current peaks, use less power during acceleration, and always accelerate the motor in the same time regardless of variations in load. Current-limit starters are desirable for motors driving high-inertia loads. A typical current-limit starter is shown in Fig. 20-60. The relays *SR*1-*SR*2 have normally closed contacts connected in series with the coils of the accelerating contactors. The coils of these relays are connected in the main motor circuit. The relays are provided with an adjustment so they can be set to close on a selected value of current.

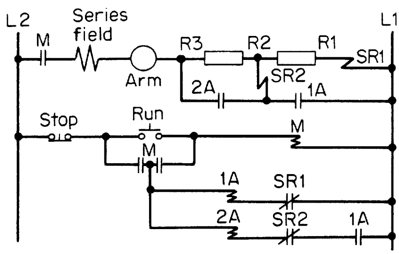


FIGURE 20-60 Direct-current series-relay accelerations.

Pressing the "run" button energizes the main contactor *M*, which closes and connects the motor to the line in series with the *R*1-*R*3 resistor. Motor current will flow through the *SR*1 coil, and its contacts will open rapidly and prevent 1*A* from closing. When the motor has accelerated enough to bring the line current down to the value for which *SR*1 is set, the relay contacts will close; a circuit is then provided for 1*A*, which closes, cutting out the first step of resistance and short-circuiting the *SR*1 coil. Current now flows through the *SR*2 coil and the 1*A* contacts. *SR*2 relay contacts open and prevent 2*A* from being energized. The motor accelerates again, and when the current falls to the value for which *SR*2 is set, its contacts close, energizing 2*A* and connecting the motor across the line.

Magnetic controllers for large dc motors are manufactured in forms to suit the application. The controllers are available in the following forms: (1) nonreversing, without and with dynamic braking; (2) nonreversing with speed regulation by field control, without and with dynamic braking; and (3) reversing with dynamic braking, without and with speed regulation by field control.

20.9.3 Synchronous Motor Starting

Methods of Starting Synchronous Motors. The method used to start a synchronous motor depends upon two factors: the required torque to start the load and the maximum starting current permitted from the line. Basically, the motor is started by using the damper windings to develop asynchronous torque, or by using an auxiliary motor to bring the unloaded motor up to synchronous speed. Recently, solid-state frequency converters have been designed to bring up to speed large several-hundred-MVA synchronous motor/generators for pumped-storage plants.

Synchronous-motor starters of the full-voltage type connect the motor directly to the supply lines. The field winding is short-circuited through a discharge resistor during the starting period. The field is connected to the dc lines when the motor is at a speed near synchronism. Reduced-voltage starters connect the motor to a reduced voltage for starting and transfer to full voltage at a speed just below synchronism. This transfer may be controlled by a time relay or a frequency relay. The field is energized either immediately before or immediately after the full-voltage switch closes. Most modern synchronous motors obtain their field voltage from a brushless exciter on the shaft.

Figure 20-61 is a simplified diagram of a synchronous motor controller arranged for starting the motor directly across the line and synchronizing by a relay operating at a selected frequency. Pressing the start button energizes relay 1*CR* and contactor *M* to connect the motor stator windings to the ac lines. Current at line frequency is induced in the rotor field winding that flows through the discharge resistor *FD* and the coil of relay *FR*. A small portion of this current flows through the reactor *X*, but the amount is limited, as the frequency is high. Relay *FR* closes rapidly, and the contacts on *FR* open the circuit to the *FS* field contactor. As the motor accelerates, the frequency of the induced current in the field winding decreases and an increasing portion of the current flows through the reactor *X*. At a speed close to synchronism most of the current flows through *X*, and there will no longer be enough current flowing through the coil *FR* to keep the relay armature closed. Relay *FR* then opens, and the contacts on *FR* close to energize field contactor *FS*. Contactor *FS* connects the field to the dc lines and opens the field discharge circuit through resistor *FD*, and the motor pulls into synchronism.

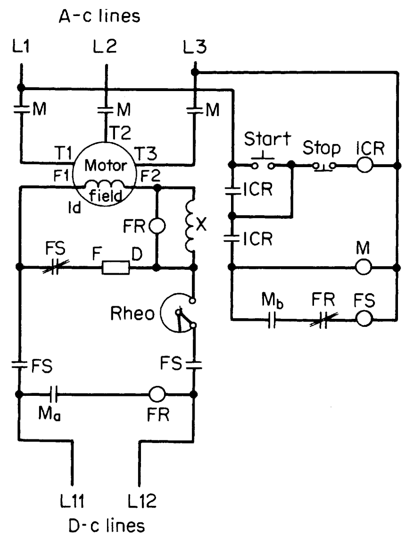


FIGURE 20-61 Full-voltage synchronous motor controller with synchronization based on frequency.

Relay *FR* is polarized by a coil connected across the dc lines through interlock contacts M_a. Polarizing the synchronizing relay provides a means for energizing the field contactor at a point in the ac wave most favorable to synchronism. Brushless synchronous motors use electronic circuits on the rotating portion to control the switching time for the field.

Speed-Control Devices. Speed control of electric motors may be obtained by various means. The design of a speed-regulating controller is determined by the type of motor with which it will be used. Table 20-17 lists the various types of motor in general use and the corresponding type of speed control for each.

Multispeed squirrel-cage motors are suitable for applications that require up to four operating speeds but that do not require speed control between these fixed speeds. However, a solid-state inverter plus a single-speed motor might be less costly than a multispeed motor, and provide much more cost-efficient operation.

Controllers for multispeed ac squirrel-cage motors may be either the drum type or the magnetic type. Drum controllers are widely used, because the many changes in connections required to obtain different speeds can be readily accomplished. Drum controllers can be used with reconnected winding or separate-winding-type motors and with constant-torque, variable-torque, or constant-horsepower motors. Low-voltage and overload protection can be obtained by using a magnetic contactor and overload relay. When complete control by push buttons or other pilot devices is required, magnetic contactors are used to change the motor connections. Controllers of this type can be arranged to permit starting at any speed or to permit starting only at the slowest speed and changing to each higher speed in sequence.

Primary Voltage Control. AC squirrel-cage motors are inherently constant-speed motors when supplied directly from utility lines. Narrow-speed-range control is obtained by adjusting the primary voltage on Design D motors using saturable reactors or solid-state phase-controlled thyristors in the stator circuits. Wide-speed-range control is obtained by adjusting the primary frequency and voltage on Design B motors using motor-alternator sets or solid-state frequency converters. The frequency of 60-Hz motors is typically adjusted from 3 to 120 Hz. From 3 to 60 Hz, the voltage is raised proportional to frequency so that the motor can deliver its full rated and breakdown torque. From 60 to 120 Hz, the voltage is kept constant so that the motor can deliver its rated horsepower.

Speed is controlled with thyristors in each of the lines to the stator of the induction motor as shown in Fig. 20-62*a*. Retarding the firing angles of the thyristors reduces the stator voltage of the

TABLE 20-17 Types of Motor Speed Control

Type of motor	Type of speed control	Speed-control range	Speed drops: no load to full load, % base speed
AC, squirrel	Multispeed Pole changing by multiple windings or reconnectable single winding	Up to 4 initial speeds	Up to 5% (slip)
	Solid-state primary voltage control NEMA Design D motor	5 to 1	20%
	Stator frequency control, constant V/Hz Solid-state inverter	20 to 1 at constant torque, plus 3 to 1 at constant hp	3%
DC, shunt-wound	Stator frequency control M-G set or solid-state frequency converter	20 to 1 at constant torque, plus 3 to 1 at constant hp	Zero
AC, slip-ring	Secondary resistors connected to slip rings	3 to 1	3%, full speed; 50%, minimum speed
	Pumpback of slip-ring power M-G set or solid-state converter	20 to 1	3%
DC, shunt-wounded	Adjustable armature voltage Solid-state converter or M-G set (Ward Leonard) plus field weakening	20 to 1 at constant torque, plus 3 to 1 at constant hp	Up to 5%
DC, series-wound	Series resistors	20 to 1	Up to 100%
	Solid-state dc chopper	20 to 1	Up to 3%

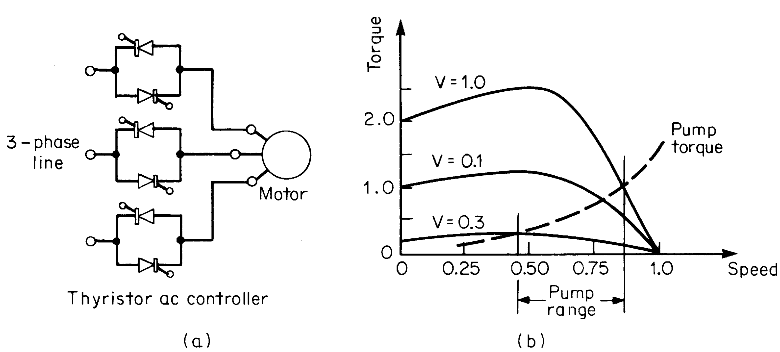


FIGURE 20-62 Primary-voltage control: (*a*) circuit of controller; (*b*) torque-speed characteristic at three-stator voltages; pump characteristic and range for 10% to 100% power.

motor. The torque at each speed is reduced as V_2, as shown in Fig. 20-62*b*. NEMA Design D motors ensure a sufficient range of descending torque in which the motor can stably drive its load. The power loss in the rotor is proportional to the torque × slip. With pump and fan loads, as shown in Fig. 20-62*b*, the torque is reduced as speed, so that the rotor power loss is acceptable at reduced speed. Typical ranges of pump and fan operation are 50% to 100% speed, 10% to 100% power.

Other Speed-, Frequency-, or Voltage-Controlling Devices. Adjustable-frequency ac induction motor drives consist of a solid-state rectifier, a solid-state inverter, the motor, and necessary controls. As shown in Fig. 20-62*a*, the rectifier converts power at 60-Hz line frequency to dc power; the inverter converts dc power to power for the motor which is adjustable in frequency and voltage. Inverters are classified by their output; they include six-step voltage, current source, and pulse-width-modulated voltage. The six-step inverter in Fig. 20-63*a* causes the motor current to approximate a sine wave.

These ac drives operate in two modes with respect to base speed, as shown in Fig. 20-63*b*. From near zero to base speed, the inverter frequency and voltage are raised in proportion so that they both reach rated value for the motor at base speed. This is termed the *constant-torque mode* because the motor can deliver its rated torque anywhere in the speed ranges below base speed. From base speed to 200% or more of base speed, the inverter frequency is raised, but the voltage is maintained constant at the rated motor voltage. The consequence is that the magnetic field in the air gap of the motor decreases and the motor is able to deliver only $^1/_7$ times its rated torque. However, the product of torque and speed is constant; the operation is termed the *constant horsepower mode*. The maximum speed depends on the mechanical capability of the motor to run above base speed and the maximum design frequency for the inverter.

AC synchronous motors are speed-controlled in special applications where their self-excitation simplifies the frequency conversion equipment and where two or more motors must operate in synchronism with the supply. Adjustable-frequency, adjustable-voltage power is supplied from a motor-alternator set, or from a solid-state frequency converter. The solid-state converter is of the cycloconverter type for speeds up to 30% of the rated speed. For wide speed range, the frequency converter is a forced-commutated inverter or a load-commutated inverter, which relies on the self-excitation of the motor to commutate the thyristors.

Synchronous motors with permanent-magnet fields rated up to about 30 hp are used in textile applications, where multiple motors are operated in synchrony from a common power supply. Synchronous motors in the range of 750 to 5000 hp and larger are used for driving large blowers, boiler fans in power plants, large pumps, and other applications requiring large motors. The block diagram for a load-commutated inverter drive is shown in Fig. 20-64. The two converters are phase-controlled thyristor bridges which can rectify or invert power, as necessary. The thyristors in the inverter operate in the firing-angle range from 90° to less than 180°, and they are commutated by the voltage generated in the stator windings of the motor by the rotating field.

AC slip-ring motor control requires that power be extracted from the rotor windings via the slip rings to reduce the motor speed, that is, increase the slip. Three methods are used: (1) secondary resistors, (2) rotor-power recovery by auxiliary rotating machines, and (3) rotor-power recovery by auxiliary solid-state rectifier and converter. The auxiliary systems recover the electric energy that would be dissipated in the secondary resistors.

AC slip-ring motor secondary-resistor speed regulators consist of a contactor to connect the primary of the motor to the supply lines and some form of resistance-switching device for the secondary circuit. The switching device may be a three-arm faceplate controller, a drum, or magnetic contactors. Regulating devices differ from starting devices in that the switching means can remain continuously on any one of the resistor steps. The motor will therefore operate continuously at a reduced speed, as determined by the amount of resistance remaining in the motor circuit. The use of secondary resistance for speed control is not an efficient method because of the power loss in the resistor. The amount of speed reduction obtained will vary directly with the load on the motor. Speed controllers of this type are usually designed for 50% speed reduction. Under favorable conditions, however, motors can be operated at 75% speed reduction.

The resistors are of the same type as the resistors used for armature regulation of dc motors.

Rotor-power recovery drives are often classified as either constant horsepower or constant torque; the designation refers to the inherent limitation in power based on full current and flux in the main

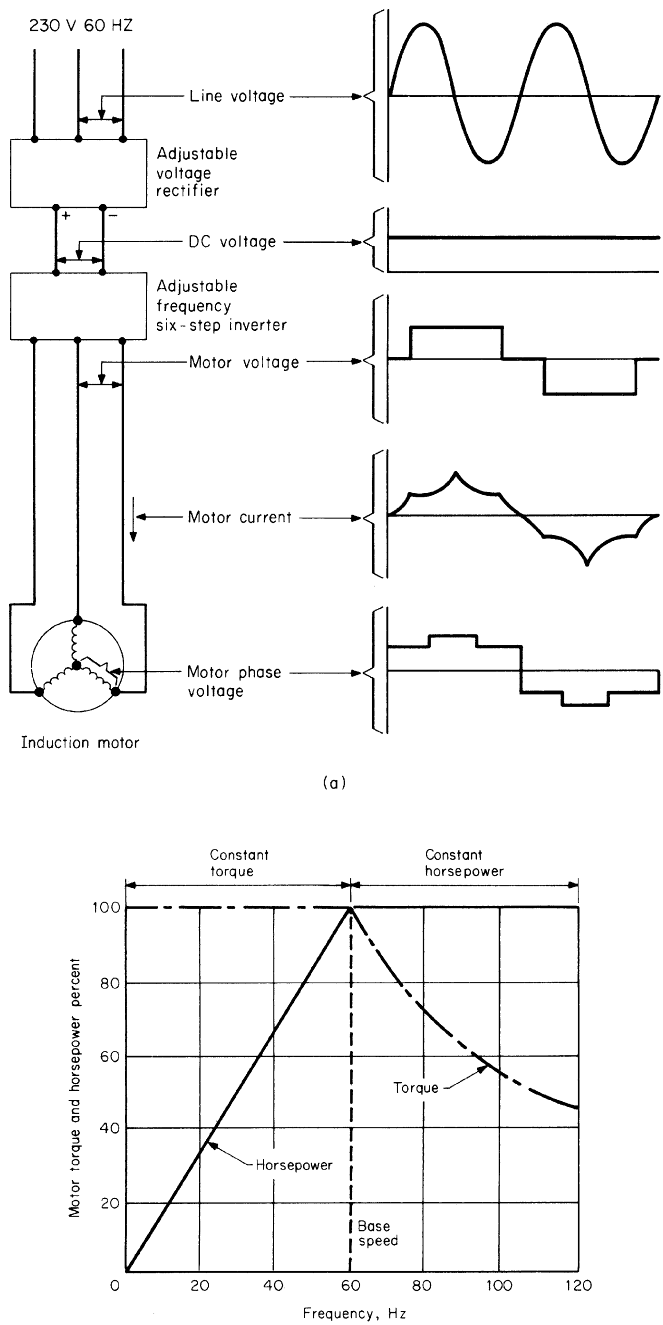


FIGURE 20-63 (*a*) Adjustable-frequency induction-motor drive using six-step inverter; (*b*) torque and horsepower capabilities of adjustable-speed drive below and above base speed.

machine. In the first scheme (constant-horsepower drive), the slip energy is converted into mechanical power and then returned to the main motor shaft. Since horsepower is a function of the product of torque and speed, such motors have high torque at low speeds and lower torque at higher speeds. In drives using this arrangement, the auxiliary machine is mounted on or mechanically geared to the main motor shaft (Fig. 20-65*a*).

In the second scheme (constant-torque drive), the slip energy is converted into electric power of the frequency and voltage of the supply circuit and is returned or fed back into the line. Since this power is not delivered to the main motor shaft, the auxiliary machine is not mechanically attached to the shaft but is separately driven. As the limiting torque of the main motor is constant, the maximum horsepower output is proportional to the operating speed (Fig. 20-65*b*).

The classical recovery systems using auxiliary machines are termed the *Scherbius drive* and the *Krämer drive.* The former employs ac commutator machines; the latter relies on a dc link and rotary converters. A variation of the Krämer drive uses a synchronous motor and a dc generator in place of the rotary converter and a constant-speed set feeding the slip power back into the line. This drive has been used for a number of large wind-tunnel drives. It is particularly adapted to a wide range of speed control and to minimum disturbance on starting.

In a solid-state version of the modified Krämer drive, the slip-ring power is rectified in a diode bridge, then passed over a dc link to a line-commutated inverter that returns the power to the supply line. The drive operates with characteristics similar to the adjustable-armature-voltage dc system. The speed is controlled by the firing angle of the inverter. The recovery systems are economical for narrow speed ranges, such as for fans and pumps, or where the horsepower rating is so large that the costs of the controls are minimized.

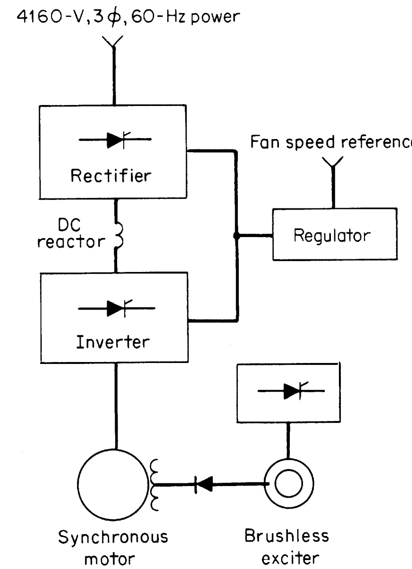


FIGURE 20-64 Block diagram of dual-converter synchronous-motor fan drive.

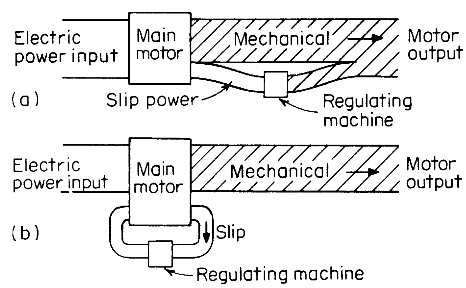


FIGURE 20-65 (*a*) Constant-horsepower drive; regulating machine, coupled to main motor, returns power mechanically; (*b*) constant-torque drive; regulating machine, mechanically separate from main motor, returns slip power electrically.

DC solid-state adjustable-voltage control utilizes phase-controlled 1-phase or 3-phase thyristor bridge rectifiers to provide armature and field voltage to independent-field dc motors. A typical armature-voltage supply is shown in Fig. 20-66*a*. As the firing angles of the thyristors are adjusted from 90° to 0°, the dc armature voltage rises from zero to rated value, and the speed follows in proportion.

The speed-torque characteristics of the dc drive are shown in Fig. 20-66*b*. In the range from zero to base speed, the field current is at maximum, but the armature voltage is raised until it reaches rated value at base speed. At any one setting of armature voltage, the speed is relatively constant over the load torque range. The firing angles of the thyristor are adjusted by a regulator to maintain the speed constant at the set point. At any speed, the rated torque is available in this range. Above base speed, the armature voltage is held at rated, but the field current is reduced to obtain the higher value of speed. The maximum output of the motor is its rated horsepower in this range.

DC series-wound motors are used for hoisting and electric traction applications where the speed is controlled either by series switched resistors or by a solid-state chopper. Where full voltage is applied, the motor operates on a varying speed-torque characteristic. For example, the speed will vary over a 10 to 1 range, as the motor load is reduced from rated to minimum torque. The series-motor controller shown in Fig. 20-67 shifts the series-motor characteristic in four resistance steps, *R*1 through *R*4. Adding resistance forces the motor to run at a lower speed for the same torque. For each resistance step, the motor has a varying speed-torque characteristic. In electric traction applications, higher motor speed is obtained by shunting the series field with additional resistors. The series motor will regenerate energy to the dc line when the series field current is supplied from another source to force the generated armature voltage higher than the line voltage.

Solid-state chopper control is replacing resistor speed control of dc series motors in rapid-transit cars operating from wayside dc lines and in battery-powered electrical industrial trucks. The chopper is more energy efficient than resistor controllers and permits stepless control of motor speed or torque. Choppers for rapid-transit cars are rated up to 1000 kW at 1000 V dc, for four-series traction motors.

A simplified circuit for a chopper is shown in Fig. 20-68*a*. The chopper applies voltage to the motor in a series of pulses, as shown in Fig. 20-68*b*. The pulse frequency is typically in the 200 to 400 pulses per second range. The width of the pulse can be varied over about a 10:1 range in a system called time-ratio control. The widest pulse provides full voltage and maximum motor speed, as shown in Fig. 20-68*c*; the narrowest pulse provides minimum voltage and speed. Each pulse is ended by turning off the thyristor with the commutating switch, which is actually another thyristor circuit. When the thyristor is turned off, the motor current continues through the diode, until the thyristor is turned on to start the next pulse. The energy loss in the thyristor is small compared to the resistors of the controller shown in Fig. 20-67. The thyristor, diode, and motor can be reconfigured for the motor to regenerate to the dc line or the battery.

Armature-regulating and field-regulating resistors are used for the speed control of dc motors fed from fixed-voltage dc supplies as shown in Fig. 20-69. Armature-regulating resistors provide starting and speed-control duty in multiple steps, usually for fan loads. The resistors are usually designated for continuous duty, in accordance

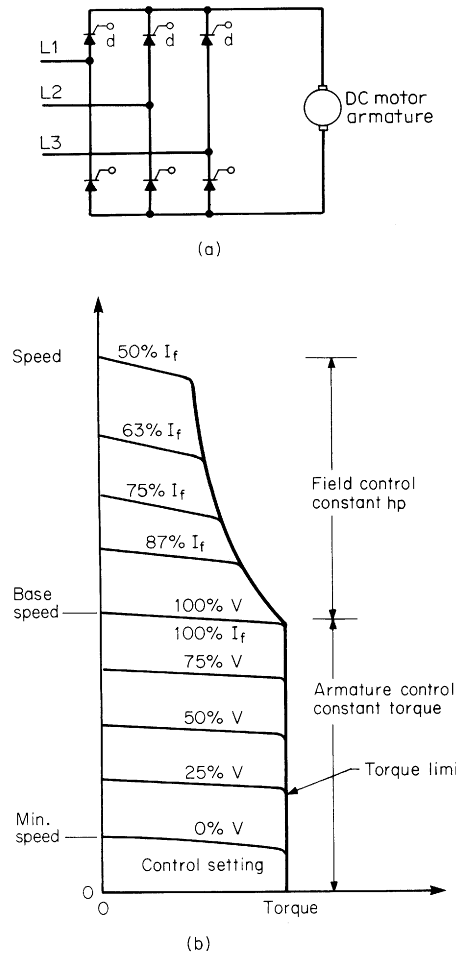


FIGURE 20-66 (*a*) 3-Phase thyristor converter for armature-voltage supply; (*b*) speed-torque characteristics of dc adjustable-speed drive.

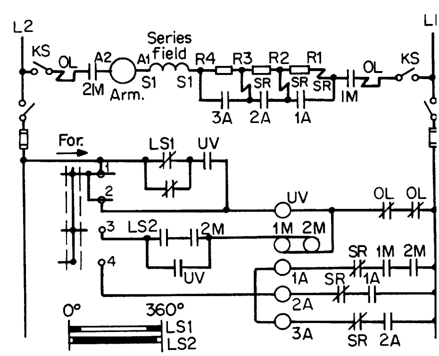


FIGURE 20-67 Series-motor controller.

with a NEMA classification number between 91 and 96, inclusive. The specific classification within this group is determined by the motor-starting current allowed, with all resistance in the circuit. The total ohms in the resistor are usually determined by the speed reduction required and the torque of the load. The *IR* drop across the resistor is the difference between the dc line voltage and the armature voltage, which is proportional to the speed. The current *I* is proportional to the torque. For example, at 50% speed a fan load requires 25% to 40% of its current at full speed. The total ohms are divided into steps to give equal speed changes per step.

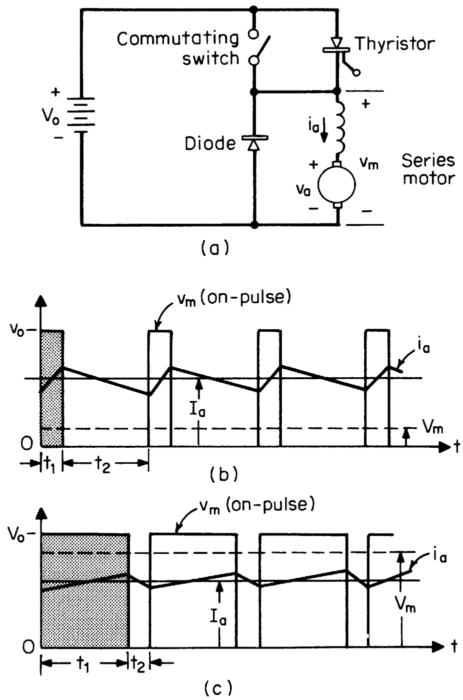


FIGURE 20-68 Chopper speed control for series motor: (*a*) circuit; (*b*) waveforms at low speed; (*c*) waveforms at high speed.

Field-regulating resistors should be designed on the basis of actual test data on the motor with which the control is to be used. The motor manufacturer will supply such data as the resistance of the shunt field or the maximum field amperes, the required rheostat resistance or the minimum field amperes. The total ohms in the rheostat will therefore be determined by the motor data. The number of steps in the rheostat must be large enough to limit the speed increase per step to an acceptable value. The character of the resistor material used may require a large number of steps, in order to keep the wattage per step on the rheostat within limits. The number of steps must be sufficiently large to keep the voltage drop at any step within specified limits to prevent arcing on the contacts. The arcing limits for sliding contacts vary from 200 V at 0.4 A to 50 V at 1 A and 25 V at 4 A. This voltage is the product of the step ohms and the amperes flowing before the step is inserted. The amount of resistor material required in the field rheostat is determined by the wattage which must be dissipated. Each step must be designed on the basis of the maximum current which it will have to carry. The total wattage capacity of the rheostat will be the summation of the step wattage. The distribution of the resistor between the various steps can be determined on the basis of obtaining either equal speed changes per step or equal percentage speed changes per step. Common practice at present is to provide equal speed changes per step, and a speed reduction to 20% of what is rated can be obtained with armature-regulating resistors. A speed increase up to 200% can be obtained by field resistors with a properly designed motor.

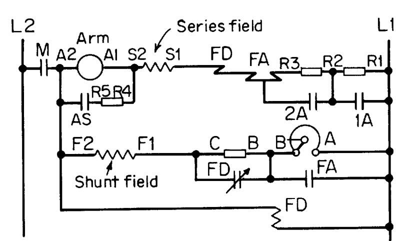


FIGURE 20-69 Armature speed-control resistors *R*1–*R*2 and *R*2–*R*3. Field speed-control resistors *A*–*B* and *B*–*C*.

20.10 STOPPING DEVICES

Electric braking is used when the dc or ac motor is allowed to make several revolutions before coming to rest. In electric dynamic braking, the kinetic energy of the motor and the load is absorbed in the resistance of the motor windings, or in a resistor switched into the circuit. In electric regenerative braking, the energy is returned to the supply line. Magnetic brakes are used to obtain quick, accurate stopping

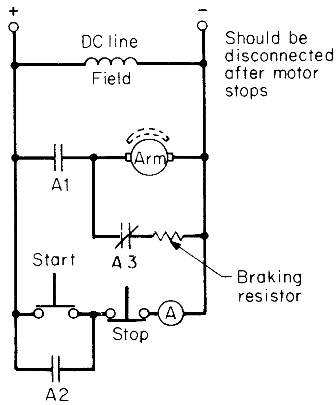


FIGURE 20-70 Dynamic braking circuit for shunt-wound dc motor.

and to hold the load after stopping. Most brakes are spring-set and electrically released, so that braking will be obtained even though an electrical failure occurs.

Dynamic braking of a dc shunt-wound motor can be obtained by disconnecting the motor from the line and shunting the armature with a resistor. As shown in Fig. 20-70, when the motor is running, contacts *A*1 and *A*2 are closed, and *A*3 is open. When the stop button is pressed, contact *A*1 opens to disconnect the armature from the line; holding contact *A*2 opens; and contact *A*3 closes to put the braking resistor across the armature terminals. The field must remain connected to the line during braking. For constant-speed motors, the ohmic value of the braking resistor is $R = (E - I_a R_a)/I$. The value of *I* determines the amount of braking obtained and may vary from 150% to 300% of normal current. With shunt motors, having speed regulation by shunt-field control, the general practice is to strengthen the field during braking by short-circuiting the field rheostat. When this is done, the ohms in the braking resistor should be $R = (E - I_a R_a)/I \times$ speed range by field. The braking obtained decreases as the motor speed is reduced. The final stopping of the motor is due to friction, since at standstill no braking torque is obtained. For very quick stopping the braking resistance can be reduced in several steps as the speed decreases, thereby keeping the current at a high value. This type of braking is effective where it is necessary to stop a motor quickly. Mechanical brakes must be used to hold a load at a standstill.

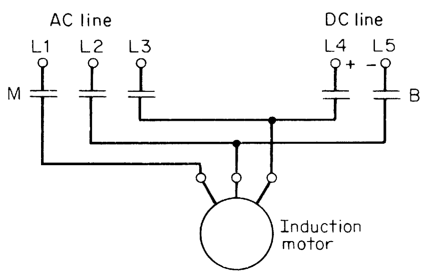


FIGURE 20-71 Dynamic braking circuit for 3-phase induction motor. Contacts *M* are for line-starting the motor. Contacts *B* are for applying direct current for braking.

Dynamic braking with an ac induction motor can be obtained by disconnecting the motor from the power supply and applying direct current to one of the stator phases. As shown in Fig. 20-71, the kinetic energy of the rotor and the load is dissipated in the rotor circuit resistance.

Plugging is used with *dc motors* to obtain very rapid reversing and is accomplished by connecting the motor to the line in the reverse direction while it is still rotating in the forward direction. The countervoltage of the armature is added to line voltage to force current through the armature and series resistor. The total resistance in circuit for plugging should be $R_t = 2(E - I_a R_a)$/inrush current. With the armature drop assumed as 10%, the formula may be written $R_t = 1.8E/1.5I = 1.2E/I$, where E = line voltage and I = normal current. The plugging resistance is determined by subtracting the armature resistance from R_t.

Plugging is used with *ac motors* to obtain a very quick stop. Three-phase squirrel-cage and wound-rotor motors may be plugged by reversing the line connections to any two of the stator terminals while the motor is running in the forward direction. In order to use plugging as a stopping means, a zero-speed switch is necessary to open the reverse contactor and prevent reversing the motor. A common form of zero-speed switch is the friction type, in which a contact is held closely by the friction of a small belt over a pulley driven by the motor. Any slight reversal of the motor will cause the contacts to open. Another type of zero-speed switch uses a disk which rotates in the magnetic field produced by Alnico magnets. Eddy currents are induced in the driven disk, and the magnetic reaction turns the magnet assembly to close the contacts. The contacts can be adjusted to open when the motor is near zero speed.

An electrical plugging relay is also used which remains closed until the line current decreases to the normal inrush current corresponding to the starting condition, or zero speed on the motor. This device has the advantage of not requiring a mechanical connection to the motor or machine.

Regenerative braking is used with dc adjustable-armature-voltage controllers to obtain rapid stopping. For a solid-state armature converter, either the dc motor field current is reversed to reverse the armature voltage, or a second reverse polarity converter is switched in. In the first case, the direction of the armature current remains the same, but the armature voltage reverses. In the second case, the direction of the armature current reverses, but the armature voltage remains the same. The regulator controls the firing angles of the converters to maintain prescribed armature current during the braking period. For an ac-motor-driven dc generator supplying the dc motor in a Ward Leonard system, the generator voltage exceeds the countervoltage of the driven motor, and power is taken from the ac lines to keep the driven machine rotating. If the field strength of the generator is decreased, the generator voltage becomes less than the countervoltage of the motor and the motor feeds power back to the generator and to the ac lines. When the countervoltage exceeds the generator voltage, a heavy reverse current is obtained, as the value of this current is limited only by the low resistance of the loop circuit. A very rapid stop is obtained, since the voltage across the generator field can be reduced to zero in about 3 s.

Direct-current brakes are set by a spring and are released by means of a solenoid or a direct-operating magnet. The coil in the operating device may be for either series or shunt connection and for continuous or intermittent duty. Series-wound brakes are operated by motor current and require 80% full motor current to release with a continuous-duty coil and about 40% full-load current with an intermittent-duty coil. The brake will be held released on about 10% full-load motor current. Intermittent-duty series brakes are rated as either $^1/_2$-h duty or 1-h duty, to correspond to the rating of intermittent-duty series motors. A series-brake coil will carry full motor current continuously, for a period corresponding to its rating, without exceeding a temperature rise of 75°C. Shunt brakes may be for either continuous or intermittent duty. Intermittent duty is defined as 1 min on and 1 min off or the equivalent, the longest time on not to exceed 1 h. Shunt brakes will release at 80% of normal voltage, when adjusted for rated torque. The larger-sized brakes use partial-voltage coils and protecting resistors. Series brakes have a heavy wire coil, which is less likely to give trouble, are faster in operation, and will set whenever the armature circuit is open. Data on a line of dc magnet-operated shoe brakes are given in Table 20-18.

Alternating-current brakes have three forms of operating mechanism: solenoid type, torque-motor type, and thrustor type. The smaller sizes of brakes are usually made in the solenoid type. On the larger sizes, a vertically mounted torque motor and antifriction ball jack provide a quiet, low-inrush-current operating means. On application of power to the motor, the rotary motion of the armature is transformed into straight-line motion through the antifriction jack. With the brake fully released, the torque motor is stalled across the line. When the motor is disconnected from the line, the spring in the brake overhauls the motor mechanism and applies the brake. Data on a line of brakes of this type are given in Table 20-19.

The thrustor-type operating mechanism consists of a self-contained motor-driven centrifugal pump, oil chamber, and a piston which produces a straight-line movement to release the brake.

The *brake size* for most applications can be determined by using the formula $T = 5252 \times \text{hp/r/min}$, where T = full-load motor torque in pound-feet, hp = motor horsepower, r/min = speed of shaft on which brake wheel is mounted. A brake should be selected with a torque rating equal to or greater than the full-load motor torque T. In some cases, the braking torque is determined by extreme operating conditions against which the brake must hold, for example, heavy ice loads on bascule bridges or conditions

TABLE 20-18 Direct-Current Magnet-Operated Shoe Brakes

Wheel diameter, in	Maximum torque, lb · ft: Shunt-wound, Int. duty	Shunt-wound, Cont. duty	Series-wound, $^1/_2$-h duty	Series-wound, 1-h duty
8	100	75	100	65
10	200	150	200	130
13	550	400	550	365
16	1000	750	1000	650
19	2000	1500	2000	1300
23	4000	3000	4000	2600

Note: Int. = intermittent; Cont. = continuous; 1 in = 25.4 mm; 1 lb · ft = 1.356 N · m.

TABLE 20-19 Alternating-Current Torque-Motor-Operated Brakes

Wheel diameter, in	Maximum torque, lb · ft		Voltamperes		WR^2 of wheel	Safe maximum rpm	Weight of brake, lb
	Int. duty	Cont. duty	Int. duty	Cont. duty			
10	160	125	160	105	3.1	2015	150
13	400	325	210	140	12	1550	240
16	800	600	300	240	25	1260	370
20	1600	1200	1000	470	75	1012	750
25	3200	2400	1500	550	220	806	1210

Note: 1 in = 25.4 mm; 1 lb = 0.4536 kg; 1 lb · ft = 1.356 N · m.

of unbalance on skip hoists. In these cases, the maximum load must be calculated and translated into pound-feet torque at the shaft on which the brake is mounted. Sufficient lining area is provided on all sizes of brake for the average application. However, a careful check as to lining area must be made when brakes are used for frequent stopping or for stopping high-inertia loads.

20.11 MOTOR-PROTECTING DEVICES

Fuses should be provided for motor circuits, in accordance with the NEC. The current rating of the fuse must be considerably higher than the current rating of the motor, or the fuse will blow when the motor is started. For that reason fuses do not provide adequate overload protection. They furnish protection for the motor only in case of a short circuit or a very heavy overload. Their primary purpose is to protect the circuit rather than the motor.

Magnetic-type overload relays are operated by direct magnetic action of the motor current on a plunger. The relay consists of a series coil connected in the motor circuit and a plunger which is pulled up into the center of the coil when a certain value of current has been reached. When the plunger is lifted, a contact is tripped, opening the motor contactor-coil circuit and disconnecting the motor from the line. The tripping current can be varied by adjusting the initial position of the plunger with respect to the coil. Time delay in tripping is obtained by attaching a small oil dashpot to the plunger. The time delay can be adjusted so that the overload will not trip on the starting-current inrush but will trip on small sustained overloads.

Thermal overload relays are available in the bimetallic type and the fusible-alloy type. The bimetallic type has two heaters in series with the circuit to be protected, and above these heaters are two strips of bimetallic material, which act as latches for the contact members. Bending of the bimetallic strips under heating of overload current will release the latches and allow the contacts to open. The fusible-alloy type has two heaters, each surrounding a thermal element consisting of a small tube, inside which is a loose-fitting shaft. The tube and shaft are rigidly joined by a special low-melting eutectic alloy. On overload, the increased current drawn melts the alloy, allowing the shaft to turn and the contacts to open.

Characteristics of a typical thermal overload are shown in Fig. 20-72. An inspection of these curves shows that the thermal overload adequately protects the wiring, that the fuse blows first on short-circuit current, and that the thermal relay allows the motor ample time to accelerate. A thermal overload has a

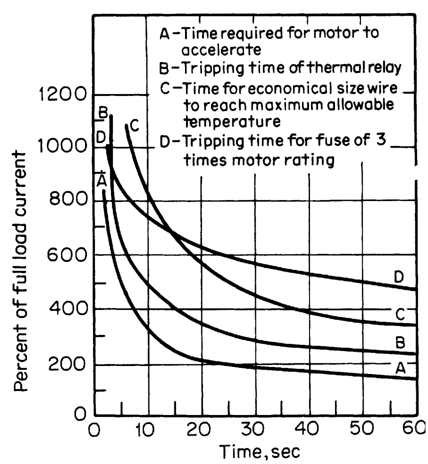


FIGURE 20-72 Characteristics of thermal overload relays.

tripping characteristic which corresponds closely to the heating characteristics of a motor and, therefore, provides an ideal protecting means. An overload coil should be selected so that the maximum permissible output can be obtained from the motor. A motor rated 40°C rise on the basis of 40°C ambient temperature will have a final safe temperature of approximately 95°C and will operate at 15% overload continuously without overheating. An overload coil should therefore be selected having an ultimate tripping current equivalent to 15% overload on the motor. A continuous overload of 15% would therefore ultimately trip the thermal relay. For overloads in excess of 15%, the tripping time would be shorter than the time required for the motor to reach a dangerous temperature.

Low-voltage protection is the effect of a device, operative on the reduction or failure of voltage, to cause and maintain the interruption of power to the main circuit.

With magnetic controllers, this protection is obtained by using some form of 3-wire master switch. Should the line voltage drop to a low value or fail altogether, the main-line contactor will open and remain open, stopping the motor. To restart, it is necessary to push the "start" button. This type of control should always be used where the unexpected restarting of a motor after voltage failure may be dangerous to workers or equipment.

Low-voltage release is the effect of a device, operative on the reduction or failure of voltage, to cause the interruption of power to the main circuit but not to prevent the reestablishment of the main circuit on return of voltage. Such protection is obtained when a 2-wire pilot device, for example, a snap switch, float switch, or pressure switch, is used.

Phase-failure protection is the effect of a device, operative upon the failure of power in one wire of a polyphase circuit, to cause and maintain the interruption of power in all the wires of the circuit.

Phase-reversal protection is the effect of a device, operative on the reversal of the phase rotation in a polyphase circuit, to cause and maintain the interruption of power in all wires of the circuit. Protection of this type is necessary on elevators, where reversing of the phases would cause the car to start in a direction opposite to that in which the operator expects it to move.

Field-failure protection is usually provided in controllers for dc shunt- and compound-wound motors. The coil of a relay is connected in series with the motor shunt field, and a normally open contact of the relay is connected in the stop circuit. If the field circuit is opened, the relay will be deenergized and the motor will be disconnected from the line. This prevents overspeeding the motor owing to an open circuit in the field. A *field protective relay* is used to insert resistance in series with the shunt field whenever the motor is not running. The coil of the relay is connected in parallel with the main switch coil, and a normally open relay contact is used to short-circuit a step of resistor in the field circuit. The resistor should be designed to reduce the voltage across the field to one-half line voltage. This reduces the field wattage to one-fourth the normal value and prevents overheating the field with the motor at standstill.

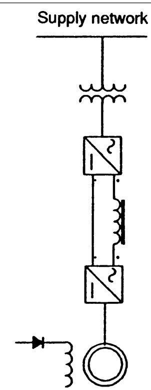
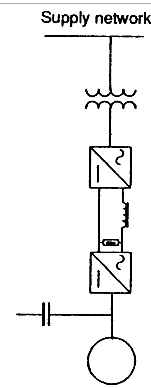
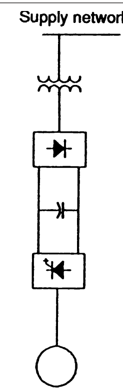
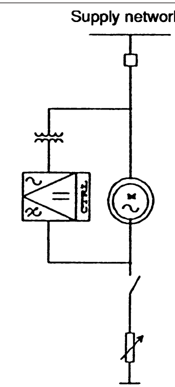
A *field-discharge resistor* should be provided for 230-V motors rated $7^1/_2$ hp or more and for 550-V motors rated 5 hp or more whenever the shunt-field circuit must be opened. The ohmic value of a discharge resistor should be 1 to 3 times the ohms in the field. If a resistance of three times the field ohms is used, the induced voltage, when the circuit is opened, will be 4 times normal line voltage. This voltage, caused by the inductance of the field, must be limited to prevent damage to the insulation of the field windings. On nonreversing controllers without dynamic braking, the shunt field can be connected behind the main contactor and the field allowed to discharge through the motor armature.

20.12 AC DRIVES

General Theory of AC Drives. There are three basic types of ac machine—synchronous machines, wound-rotor induction machines, and squirrel-cage induction machines (see Table 20-20). Each of these machines is normally applied to a fixed-speed application. The speed of each is determined by the internal mechanical and electrical configuration. It can be displayed as

$$\text{Speed (r/min)} = \frac{120 \times \text{frequency (Hz)}}{\text{No. of poles}}$$

TABLE 20-20 Basic Types of AC Machine

Drive system	LCI-fed synchronous motor	LCI-fed squirrel-cage induction motor	Voltage-source PWM inverter-fed squirrel-cage induction motor	Wound-rotor induction motor with slip-recovery converter
System one-line diagram				
Type of machine	Synchronous motor	Squirrel-cage induction motor	Squirrel-cage induction motor	Wound rotor induction motor
Typical power range	1000–100,000 hp	1000–20,000 hp	1000–10,000 hp	1000–20,000 hp
Maximum speed	7500 rpm (depending on power)	7500 rpm (depending on power)	7500 rpm (depending on power)	1800 rpm
Typical speed range	10–100%	50–100%	0–100%	50–98%
Significant properties and features	Simplest converter Single-motor drive Four-quadrant operation	Applicable to existing SCIMs Nearly sinusoidal motor voltage and current	High-power factor Good dynamic performance over entire speed range Low harmonic distortion	Suitable for retrofit to existing slip-ring motor Very economical for narrow speed ranges Inherent bypass capability

Except for special cases, the number of poles in a machine is fixed and the incoming line frequency is normally fixed. Therefore, except for slip in induction machines, a given ac motor operated on a given electrical network will operate at a single fixed speed.

If the line frequency supplying the motor can be adjusted, then the speed of the motor can also be adjusted. That is the basis for adjustable-speed drives for ac motors.

Application and Economics of Speed Control. As previously noted, ac motors provide the motive power for a wide array of industrial applications. Most of these involve a centrifugal type of load, for example, fans, blowers, compressors, or pumps. The configuration of these devices is such that the *torque* required to turn them is proportional to the square of the speed. And the *power* required to turn them is proportional to the *cube* of the speed. The relationships

$$\text{Torque} \propto (\text{r/min})^2$$

$$\text{Power} \propto (\text{r/min})^3$$

is what makes adjustable-speed drives (ASDs) economically attractive.

Most of these devices are part of a process that requires some degree of flow or pressure control. This is normally performed via mechanical throttling of the flow, for example, inlet or outlet dampers, guide vanes, valves, or bypass systems. The net result of any of these methods is the consumption of excess energy required to overcome the pressure drop or bypass flow each creates.

However, efficient and precise control of the flow or pressure in a system can be obtained by directly controlling the speed of the device. This direct-speed control, coupled with the speed/power and speed/torque relationships shown above, results in an exceptionally efficient method of controlling the process. Several programs are commercially available to calculate energy savings and paybacks based on energy rates, operating profiles, internal rates of return, etc.

Power Semiconductor Devices. In order to supply an adjustable frequency to an ac motor, the incoming constant-frequency voltage or current must first be *rectified* to a dc current or voltage. This dc voltage or current is then *inverted* to an adjustable frequency output. This is accomplished by power semiconductor devices.

These devices allow current to flow in only one direction. They act as electronic switches. Various types can be fired on and off at certain times in the cycle. This selected firing allows control of the resultant voltage or current.

Power electronic devices come in three basic types (Fig. 20-73):

Diodes. This is an uncontrolled device, that is, current can flow only in one direction. However, there is no gate to control the firing angle.

Thyristors. These devices also allow current to flow in only one direction. However, they include a gate circuit to delay firing until a given angle in the sine wave. This allows control of the resultant output dc current. However, it normally is turned off or *commutated* only when the voltage across it drops to zero.

Gate turnoff (GTO) *thyristors*. These are similar to basic thyristors. They also include a turn*off* gate that allows the current to be blocked at any point in the cycle.

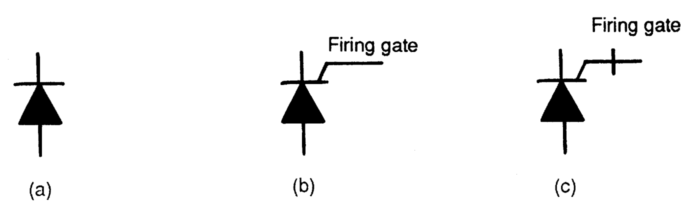


FIGURE 20-73 Main power electronic device used in modern frequency converters: (*a*) diode; (*b*) thyristor; (c) gate turnoff (GTO) thyristor.

AC Drive Systems. Electronic speed control offers much more than just energy savings. Reduced maintenance and reduced mechanical and thermal stresses are just a few of the advantages that can be realized.

The key to maximizing the many advantages of adjustable speed drives is applying the proper type of drive system. Following is an overview of the most important types of drive systems available today.

Most drive systems are composed of three major components—the input transformer, the frequency converter, and the ac machine. Additional components may include dc link reactors, harmonic filters, output filters, switchgear, etc. The requirements for these components will depend on the type of drive, user specifications, and vendor recommendation.

Connection to the Electrical Network. Most types of drive systems are connected to the electrical network through an input transformer (Fig. 20-74). This transformer serves many purposes. It transforms the network voltage to the optimal voltage for the frequency converter. It effectively lowers the short-circuit capacity seen by the drive to reduce "line notching" that can occur when thyristors commutate from one phase to another. It also provides impedance in the line to reduce distortion. A line reactor can be used if the line voltage and converter voltage match. However, a three-winding transformer can be used to dramatically reduce harmonic distortion. The two secondary windings are phase-shifted 30 electrical degrees by connecting one in a delta configuration and one in a wye configuration. This effectively eliminates fifth and seventh harmonics and their multiples.

LCI/Synchronous Motor Drives. In this system, the line-commutated rectifier acts with the dc link smoothing reactor as a dc current source. The load-commutated inverter operates with the synchronous machine's voltage and frequency to switch the dc current to the windings. The synchronous machine behaves like a dc machine with the inverter operating as a static commutator.

The load-commutated inverter (LCI) (Fig. 20-75) is the simplest type of frequency converter. Because the rectifier thyristors are commutated by the line voltage and frequency, and the inverter thyristors are commutated by the synchronous machine voltage and frequency, no special circuits are needed to accomplish commutation.

This type of drive has the added inherent advantage of 4-quadrant operation. Therefore, it is suitable for drives which may require regenerative braking or operation in both the motoring and generating modes.

LCI/Induction Motor Drives. An induction motor can also be driven with an LCI-type drive (Fig. 20-76). However, because of the inherent characteristics of an induction machine, the thyristors in the inverter section cannot be self-commutated by the motor voltage until it reaches approximately 50% speed. Therefore, additional circuitry must be added across the inverter or across the individual thyristors to force the voltage to zero and turn off the thyristors. Additionally, a filter on the output of the inverter is normally required to produce a sinusoidal current form to the motor.

This type of drive is commonly used as a retrofit to existing squirrel-cage induction motors. However, care should be exercised on such applications. All frequency converters impose additional losses and, therefore, additional heating on motors. Higher stresses are also placed on the insulation

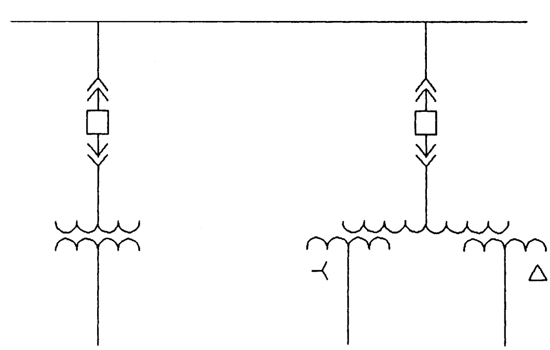


FIGURE 20-74 Two- and three-winding transformers.

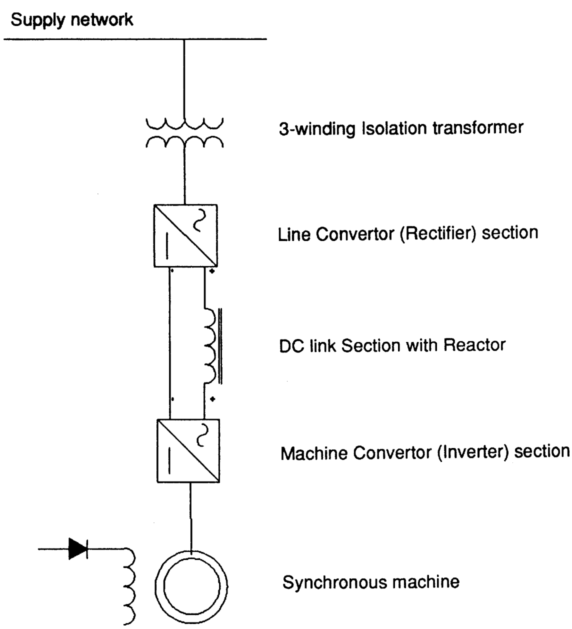


FIGURE 20-75 Basic LCI system design.

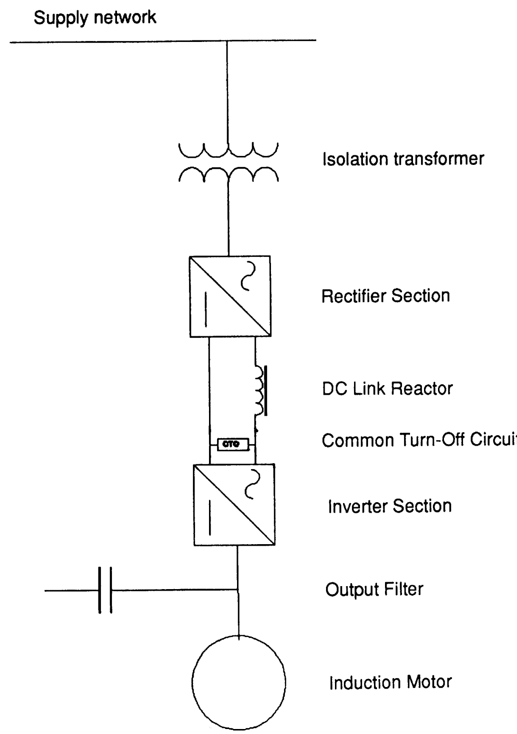


FIGURE 20-76 LCI drive with induction motor.

of the motor windings due to the distortion of the inverter's output waveform.

PWM/Induction Motor Drives. With the advent of high-power GTO thyristors, pulse-width modulated (PWM) frequency converters for powers up to 10,000 hp are now available. These can be supplied as either voltage-source (Fig. 20-77) or current-source configurations. For either configuration, the ability to *turn off* the device via a gate pulse eliminates a number of components in the power circuit.

The voltage source drive in particular has a number of benefits. The uncontrolled diode rectifier section results in a high line-side power factor (near unity) throughout the speed range. The distortion seen by the electrical network is also reduced since diodes always have an effective firing angle of zero. The PWM voltage-source output of the inverter produces a near-sinusoidal current in the motor. High dynamic performance can also be realized. Full torque can be maintained and controlled throughout the speed range—even at zero speed.

A unique application of voltage-source PWM drives is their use on coordinated sectional drive systems. Because the voltage in the dc link is maintained at a constant value, several inverter sections can be connected to the same rectifier/dc link section. This can result in significant savings in money and in space.

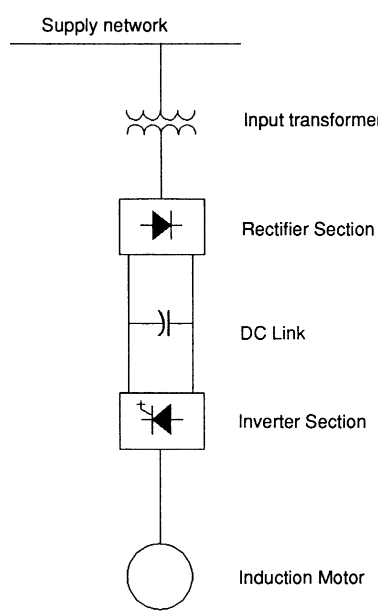


FIGURE 20-77 Voltage-source PWM inverter drive.

Slip Recovery (Cascade) Drives. A much different solution is a slip recovery, or cascade, drive system (Fig. 20-78). While the previously described systems vary the voltage and the frequency applied to the stator of the ac machine, a cascade drive is connected in the rotor circuit of a wound-rotor motor. Wound-rotor or slip-ring motors can be speed-controlled by connecting an external resistor to the rotor circuit via the slip rings. Since the slip (and hence the speed) of an induction motor is directly proportional to the losses in the rotor, increasing or decreasing the external resistance could vary the speed over a limited range. Historically, the resistance was adjusted via a liquid rheostat for stepless control, or a grid resistor for several steps in resistance. This method was also commonly used for starting of high inertia loads. The motor's starting torque and current characteristics could be adjusted during starting. Once it reaches rated speed, the slip rings are short-circuited and the motor functions as a squirrel-cage induction motor.

The cascade drive controls the speed in a similar fashion. However, as its name implies, rather than dissipating the energy across external resistors, it *recovers the slip energy* and returns it to the electrical network. This is realized by converting the low-frequency rotor current into line-frequency current. It is then transformed to the supply voltage via a feedback transformer. Starting is normally accomplished via conventional means with a liquid or grid resistor.

The cascade drive is an excellent solution when the speed range is relatively narrow—normally 50% to 70% to 100%. Because the converter handles only the *slip* power, it is much smaller than other types of frequency converters. As an example, if the desired speed range is 70% to 100%, the converter must be sized to handle only 30% of the rated power. This results in a very economical solution.

An additional advantage of this type of drive is that a bypass arrangement is inherent in the design. The normal starting sequence is to bypass the drive and connect it once the motor reaches a minimum control speed. Therefore, if the drive malfunctions for any reason, it can be bypassed and isolated. In this way repair can be accomplished without shutting down the process.

Other Drive Types. There are other types of ac drive systems used for specialized applications. One of these is the *cycloconverter* drive. This drive is normally applied to low-speed, high-torque, high-performance applications. These include metal rolling mills, large grinding mills for ore or in

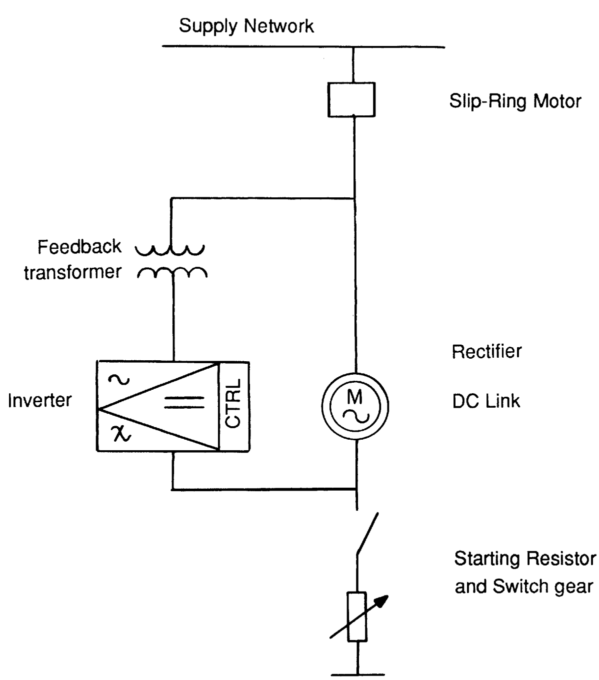


FIGURE 20-78 Slip recovery (cascade) drive.

cement applications, and ship propulsion. All these applications require precise speed and torque control to optimize the process.

The cycloconverter is effectively a separate frequency converter on each phase of a synchronous or induction motor. This allows for a more precise fabrication of the voltage waveform to the motor. This, in turn, results in a current wave shape that very closely approximates a sine wave. Consequently, the torque produced by the motor is very smooth throughout the speed range.

The design of the cycloconverter is such that its output frequency cannot exceed one-third of the line supply frequency. So, for a 60-Hz system, the maximum output frequency is 20 Hz. A 12-pole motor, which would normally have a synchronous speed of 600 r/min on a 60-Hz system, cannot be run up to only 200 r/min with a cycloconverter drive.

Another type of relatively exotic drive is the supersynchronous converter cascade. This is used for matching frequencies from a variable-frequency generator, for example, a hydropower station with varying head and speed.

Harmonic Impact of Drives on Electrical Networks. *All* drives result in some harmonic distortion in the electrical network to which it is connected. This is due to the fact that the power semiconductor devices "chop up" the incoming sine wave and reconstruct it at another frequency. By definition, this chopping up of the waveform results in distortion from a pure sine wave.

Harmonic distortion can have effects in several areas. Electronic equipment, such as computers, can malfunction due to changes in zero-crossing points in the supply waveshape. Motors may run hotter due to harmonic currents.

Many factors affect the magnitude of the distortion. The type of rectifier section on the drive has a major effect. Uncontrolled diode bridges will typically produce less distortion. Thyristor rectifier bridges will be somewhat higher since the firing angle is greater than zero.

The short-circuit capacity of the system at the point of connection is the major factor in determining distortion. The distortion can be estimated from Fig. 20-79. These curves clearly indicate that if the drive is connected to a relatively strong system, that is, one whose short-circuit capacity (SCC) is substantially higher than the converter rating, the total harmonic distortion seen by the network may not exceed the limits established by IEEE 519.

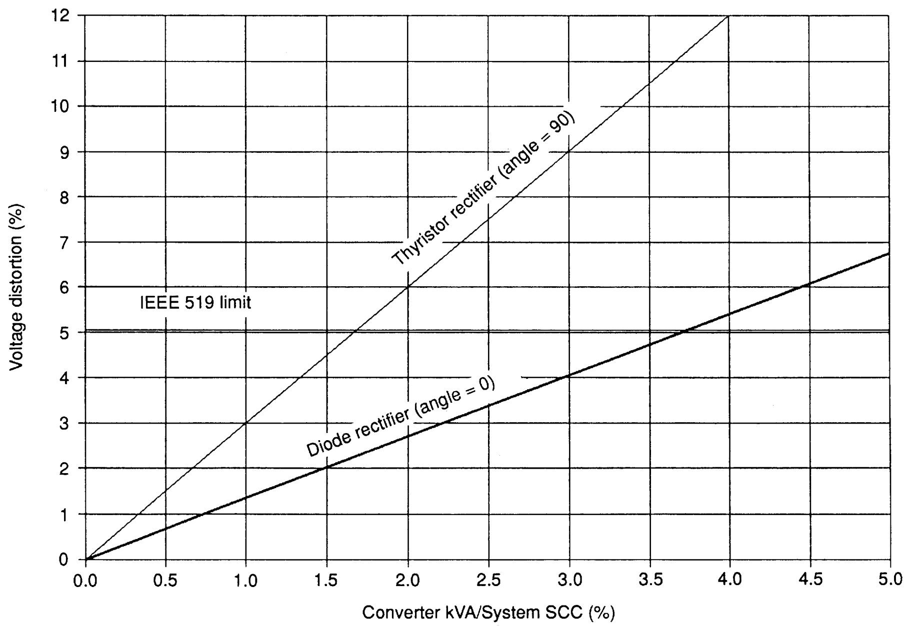
Connecting to the network with a 12-pulse configuration (through a 3-winding transformer) has a great impact on the distortion. The connection effectively eliminates the fifth and seventh harmonics and all multiples of them. Since the magnitude of the harmonic is inversely proportional to the order number, these are the greatest contributors to distortion.

If multiple drives are connected to the same bus, distortion can be further minimized by phase shifting the primaries of the drives to effectively create a 24-pulse connection. This is only feasible with drives utilizing a diode rectifier. This is due to problems with timing of the device firing on thyristor-type rectifiers.

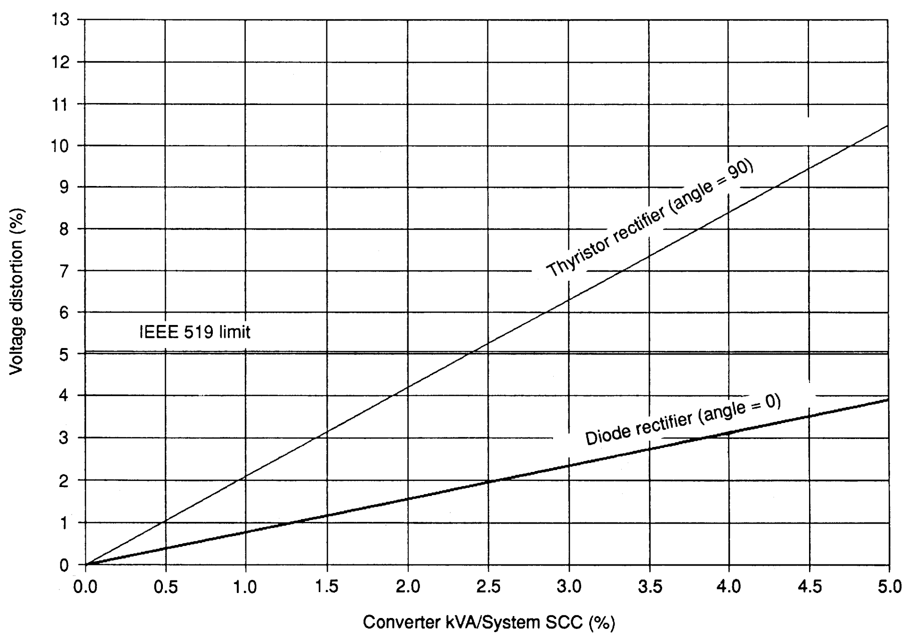
If the type and configuration of the drive do not limit the harmonic distortion to an acceptable level, filters can be installed on the bus to further reduce their effect. These filters are configured with reactors, capacitors, and resistors to "trap" the unwanted harmonics. Typically they are "tuned" to a specific order number, that is, normally there are fifth, seventh, and eleventh harmonic filters. These filters need to be properly coordinated with the drive supplier and the electrical network to ensure that resonances are not created that could interact with other equipment.

Impact of Drives on Mechanical Systems. Careful consideration should be given to the *mechanical* system when applying a drive. One key point may be the natural critical speed or resonance point of the motor. When applied as a constant-speed machine, the motor may operate above or below its natural frequency with little problem. However, if the critical speed is within the desired speed control range, problems can occur. These can be addressed in several ways. Critical speeds can be changed somewhat via bearing or base configurations. Or, the drive can be programmed to avoid certain frequency or speed bands. In this case, the drive will quickly accelerate or decelerate the motor through the resonance points to minimize potential problems.

Just as the electrical network sees distortion from a drive, the connected load (normally a motor) also sees an imperfect sine wave as the output of the drive. This distortion of the incoming current



(a)



(b)

FIGURE 20-79 Harmonic distortion: (*a*) 6-pulse rectifier; (*b*) 12-pulse rectifier.

or voltage is translated into a distortion of the motor torque. This can be realized as torque pulsations at the motor shaft output. For many applications, this does not present a problem. However, for all applications, a detailed torsional analysis of the complete mechanical system should be performed to ensure that the drive will not excite any torsional resonance points. If such problems are uncovered early, they can usually be easily addressed through modifications to the drive or through a flexible type of coupling to dampen the torsional vibrations. The system torsional analysis may be performed by the drive vendor, the driven equipment vendor, or both.

Summary. Properly applied, ac drive systems can have tremendous positive impact on a wide range of processes and applications. Carefully engineered drive systems can provide many years of trouble-free operation.

BIBLIOGRAPHY

Books and Publications

Andreas, J. C., *Energy-Efficient Electric Motors, Selection and Application,* New York, Marcel Dekker, 1992.

Bose, B. K., *Adjustable Speed AC Drive Systems*, New York, IEEE Press, 1980.

Cochran, P. L., *Polyphase Induction Motors, Analysis, Design, Application*, New York, Marcel Dekker, 1989.

Fitzgerald, A. E., Kingsly, C., Jr., and Kusko, A., *Electric Machinery*, 3d ed., New York, McGraw-Hill, 1971.

Murphy, J. M. D., *Thyristor Control of AC Motors*, New York, Pergamon, 1973.

Nasar, S. A., and Boldea, I., *Linear Motion Electric Machines*, New York, Wiley, 1976.

Say, M. G., *Alternating Current Machines*, 5th ed, New York, Wiley, 1983.

Shoults, D. R., Rife, C. J., and Johnson, T. C., *Electric Motors in Industry*, New York, Wiley, 1942.

Smeaton, R. W., *Motor Application and Maintenance Handbook*, New York, McGraw-Hill, 1969.

Alerich, W. N., and Herman, S. L., *Electric Motor Control*, Albany, N.Y., Delmar Publishing, 1998.

Beaty, H. W., and Kirtley, J. L., Jr., *Electric Motor Handbook*, New York, McGraw-Hill, 1998.

Chapman, S. J., *Electric Machinery Fundamentals*, New York, McGraw-Hill, 1998.

Dorf, R. C. (ed.), *The Electrical Engineering Handbook*, 2d ed., Boca Raton, Fla., CRC Press, 1997.

Emadi, Ali, *Energy-Efficient Electric Motors*, New York, Marcel Dekker, 2005.

Hamdi, E. S., and Hamdi, H. S., *Design of Small Electrical Machines: Design and Measurement in Electronic Engineering*, New York, John Wiley & Sons, 1994.

Hughes, A., *Electric Motors and Drives: Fundamentals, Types and Applications*, Oxfordshire, England, Butterworth-Heinemann, 1993.

Kaiser, J., *Electrical Power: Motors, Controls, Generators, Transformers*, Tinley Park, Ill., Goodheart-Willcox, 1998.

Keljik, J., *Electrical Motors and Motor Controls*, New York, Delmar Publishing, 1995.

Novotny, D. W., and Lipo, T. A., *Vector Control and Dynamics of AC Drives*, Oxford, England, Clarendon Press (imprint of Oxford University Press), 1996.

Subrahmanyam, V., *Electric Drives: Concepts and Applications*, New York, McGraw-Hill, 1996.

Web Sites

American National Standards Institute: http://ansi.org/IEEE Press: http://shop.ieee.org/

Library of Congress: http://catalog.loc.gov/McGraw-Hill.Professional Books: http://books.mcgraw-hill.com/Wiley Publications: http://wiley.com/

SECTION 21
INDUSTRIAL AND COMMERCIAL APPLICATIONS OF ELECTRIC POWER

Hesham Shaalan
Associate Professor of Electrical Engineering, U.S. Merchant Marine Academy, Kings Point, NY

CONTENTS

Most of the original material in this Section was developed by engineers at Resource Dynamics Corporation, Vienna, VA.

21.1 INTRODUCTION

21.1.1 Links between Competitive Advantage, Efficiency Improvement, and Environmental Compliance

Since the last edition of this handbook, a newly competitive environment is emerging in the industrial and commercial applications of electric power. While the areas of improving energy efficiency and meeting stricter environmental regulations is still a concern to business and industry, the need to maintain a competitive edge in an increasingly global economy is having a definite impact on energy-related decisions. Technology investments are still being made in process and business enhancements, but the driving force is business economics. Producing goods and delivering services in a way that is "cheaper, better, and faster" is the goal of most competitive organizations. Technologies at the forefront of improving business operations include sophisticated information and communications systems, new sensors and control systems, and constantly improving electrotechnologies. The success of electrical engineering today will depend to a great degree on the extent to which the engineer understands this technological changes, and participates in the business decision making of the company. Nevertheless, the ultimate objective of any successful business is still to improve performance while cutting costs.

Sensors. From an electrical engineering perspective, sensors are an essential element in the operation and control of a manufacturing or other electricity-driven process. Sensors include all devices that respond to a physical, chemical, or biological stimulus and transmit a resulting impulse for measurement or control. Simple devices include electrochemical sensors that determine ionic or molecular concentration, potentiometric sensors that measure the potential difference between two electrodes, and amperometric sensors such as the Clark cell for measuring oxygen in some fluids (see Sec. 24 for more details). In all cases, the sensor typically responds to a change in condition, and converts the measured variable into an electric signal.

New advances in sensor technology encompass on-line machine diagnostics, remote and noninvasive detection, and improved durability in hostile environments. Sensors are used in industrial plants and mills to control process flows, assembly-line speeds, chemical concentration levels, and many other variables. A typical application of sensors is in industrial flexible manufacturing systems, assemblies of one or more machine tools and workpiece-handling devices, inspection sensors, and part-washing equipment and/or material storage equipment, all operating in a coordinated manner under the control of a central or distributed computer. A flexible manufacturing system is employed to process a variety of finished parts (see Fig. 21-1). Commercial building facility management systems include sensors for input data, remote-terminal units, the central processor, and human-machine interface devices. Functions typically go far beyond energy management, including not only heating, ventilation, and air conditioning (HVAC), but also fire management, security, and access control.

Information and Communications Systems. The information and communications systems in a plant or facility consist of collecting hardware, software, and input/output devices and connecting wire or cable that transmits voice and data, and then processing these data into information and knowledge for decision-making purposes. Improved communications systems are being developed based on high-speed data transfer and expanded application of voice recognition interface. This is the basis for emerging smart information systems that will take full advantage of language translation, natural-language processing, artificial intelligence, storage and processing of only useful data, and interactive computer-based training.

Electrotechnologies. For the electrical engineer, the technologies that use electricity to manufacture or transform a product are of special interest—collectively, these technologies are known as *electrotechnologies*. The industrial-commercial market continues to represent significant opportunities for electrotechnologies, ranging from process heating to metal heating, cutting, and welding. In most electrotechnologies, electromagnetic, electrochemical, and/or electrothermal effects are central parts of the process. Examples of these technologies include induction heating and melting; plasma processing; infrared, microwave, and radio-frequency processing; freeze concentration; and electroseparation. Some of these technologies, such as infrared heating, also have natural-gas-fired alternatives.

A broad set of electrotechnologies includes electric motors, used in the commercial and industrial sectors to drive pumps, fans, and compressors for a wide range of applications. These applications include HVAC applications, fluid processing, compressors to drive freeze concentration, and membrane separation. In the area of materials processing, motors furnish the power for cutting, grinding, and crushing. Finally, the raw materials and manufactured products are moved around the factory floor by motors driving conveyors, cranes, elevators, and robots.

21.1.2 Environmental Compliance

The environmental impact of the industrial and commercial applications of electric power has also rapidly become a primary concern in many industry sectors. Environmental concerns, many heightened by more stringent laws and regulations, are widespread, including such problems as the release of volatile organic compounds (VOCs) during solvent use for example, in industrial painting and curing and disposal of oil-water emulsions, toxic wastes, and other industrial effluents. Mitigation of these problems is typically addressed with the "TR3 approach"—treat, reduce, reuse, or recycle. For example, the generation of VOC emissions can be reduced (or eliminated) with the use of water-based paints or powder coatings combined with infrared drying. A parallel treatment option would include the use of solvent recovery heat pumps, or perhaps freeze concentration to separate VOCs from wastewater.

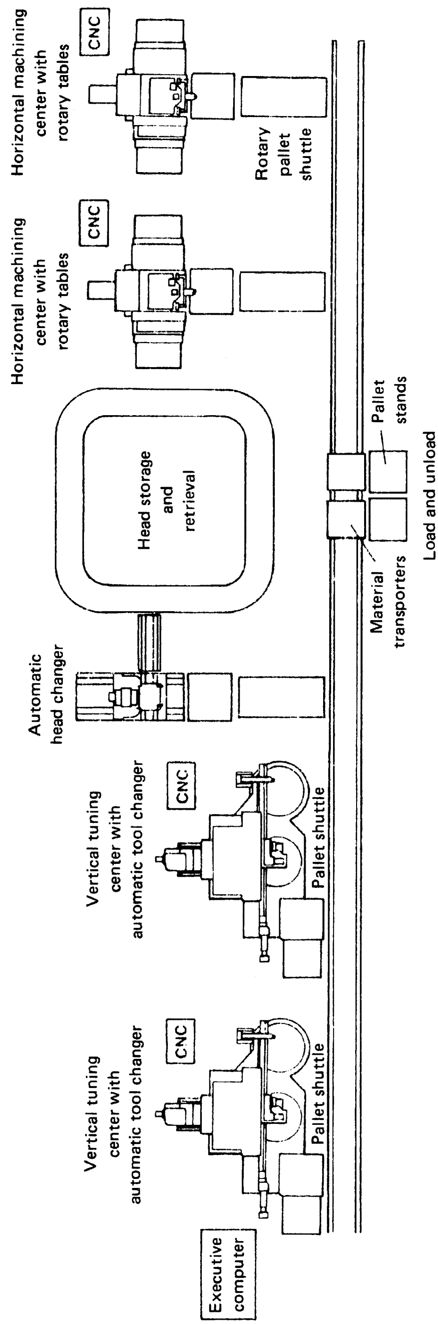


FIGURE 21-1 Schematic illustration of a flexible manufacturing system.

21.2 TRENDS IN BUSINESS AND INDUSTRY ENERGY USE

21.2.1 Impact of Deregulation

By 1900, electric utilities produced approximately two-fifths of the electricity in the United States. The balance of power was supplied by business and industry which generated their own electricity. As larger and more efficient generators and more transmission lines were installed by the electricity industry, costs came down and the associated increase in convenience prompted most customers to concentrate on their core business operations and purchase electricity from the utilities. As originally configured, the electric power industry was based on a central source of power supplied by efficient, low-cost utility generation, transmission, and distribution. Utilities were granted exclusive franchise areas in which to operate, and along with this exclusivity came the obligation to serve all consumers within that territory, and, in most cases, state regulation of privately owned electric utilities in the early 1900s. States typically regulated utility rates, financing, and service, and established utility accounting systems.

The U.S. electric power industry today is in the process of restructuring, which began with the introduction of federal deregulation policies starting with the passage of the Public Utility Regulatory Policies Act (PURPA) of 1978. PURPA was initially positioned as a way to encourage the development and use of alternative fuel resources in the industry. The important long-term effect was the introduction of some level of competition by providing new participants with a gateway to electric utility markets. The key elements of PURPA were the requirements for public utilities to purchase available power from qualified cogenerators and small power producers at rates that were less than or equal to the utility's avoided costs, and provide backup service to cogenerators and small power producers at nondiscriminatory and reasonable rates. In addition, the nonutility generators were exempted from various state and federal regulations. The entrance of a new group of suppliers demonstrated that cogenerators and small power producers represented a viable source for new supply of energy and power services.

The Energy Policy Act expanded markets further in 1992, and in 1996, the Federal Energy Regulatory Commission (FERC) issued rules for implementing open access to the transmission network and for utilities to recover the costs associated with transmission lines and other plant and equipment investments that may be "stranded" as markets become more competitive. Today, the state public utility commissions are actively studying retail competition, and some have already introduced pilot programs or have drawn up plans for restructuring. Newly formed entities, such as power marketers, brokers, and independent system operators, are emerging on the scene. The electric utility companies are being merged and acquired, promising a much different power supply picture in the future.

21.2.2 Role of the Energy Service Company

As the electric utility market has slowly begun to entertain competition, a new kind of business has emerged: the energy service company. The energy service company (ESCO) may be independent, or may be a subsidiary of an electric or gas utility company, and it typically develops, installs, and finances projects designed to improve the energy efficiency and lower operating and maintenance costs for commercial and industrial facilities. ESCOs generally assume the technical and performance risk associated with a specific project, and often act as the project developer and manager. These companies not only install and maintain the energy equipment but also measure, monitor, and verify the project's energy savings. All services provided by the ESCO are usually bundled into the project's cost and are repaid through the savings generated. Projects undertaken include high-efficiency lighting, high-efficiency heating and air conditioning, efficient motors and variable-speed drives, and centralized energy management systems. ESCOs frequently work on a performance-based contracting basis, and often the company's payments are directly linked to the amount of energy that is actually saved.

21.2.3 Retail Power-Supply Options

The ongoing transformation and ever-increasing competitive nature of the electric industry has greatly enlarged the scope and complexity of how electricity will be delivered to the customer. A broad range of power technology options are available and emerging, including fuel cells, turbines, microturbines, reciprocating engines, and a range of renewable technologies (e.g., photovoltaic and wind). Both economics and reliability are also factors being considered in the development and implementation of these technologies. Many of these technologies are "distributed," meaning that they are typically sited close to the customer load and are generally smaller in size than the megawatt-sized utility units. Technologies such as solid oxide fuel cells offer a combination of performance and flexibility which makes them an ideal supply resource that helps address regulatory, environmental, and competitive challenges in delivering essential power and energy services. By converting fuel energy directly into electricity, fuel cells provide a clean and strategically economical resource. Microturbines are also getting a lot of attention, and consist of a compressor, combustor, turbine, and generator. This technology is derived from aircraft auxiliary power systems and diesel engine turbochargers. A number of companies are currently field-testing demonstration units, and commercial deliveries started in 1999.

Some deregulation measures did not succeed as originally planned. For example, California's 1996 law deregulating the electricity market, once hailed as a model for others:

- Forced utilities to sell off much of their generating capacity
- Prohibited them from signing long-term contracts to buy supplies
- Barred increases in consumer rates until 2002

People inside and outside of the state of California wonder how such problems could happen in their state, the home of Hollywood and Silicon Valley. Even though it's a complicated issue, it mostly results from geography—the state's population and businesses (especially power-draining high-tech industries) have grown by leaps and bounds over the past decade, while no new power generation plants have been built in the state over the past decade.

Additionally, power can't be stored up and used at a later time. Supply must equal or exceed demand at the very instant that the demand is there. Due to California's lack of power generation facilities, California is obtaining power from across the Western United States and less-than-adequate a rainfall in the Pacific Northwest has resulted in less power being available from the hydroelectric plants of the Northwest.

21.3 ELECTRICITY IN AGRICULTURE

There were 1.9 million U.S. farms in 1997, compared to 5.4 million farms in 1950. Farm output has increased steadily since the mid-1980s, reflecting almost universal use of automation and improved chemical fertilizers. Food products and fiber are the primary outputs of the agricultural sector.

Increased global economic growth had driven increased demand for food and fiber worldwide. The U.S. comparative advantage in agricultural production and transportation has allowed it to capture an increasing share of the growing global demand. In 1996, U.S. agricultural exports reached $60 billion, and in 1997 exports were $57 billion and farm prices were firm. Exports slipped to $55 billion in 1998 as recession hit Asia and expanded production around the world has lowered crop prices.

21.3.1 Energy Use in Agriculture

Electricity consumption in the agricultural sector has been decreasing, reflecting farm consolidation and efficiency improvement.

21.3.2 Technology Innovation

Increased automation (discussed below) controlled by increasingly sophisticated and lower-cost computer systems has helped improve farming operations across the board. Introduction of advanced technology, such as the written-pole motor, promises improved electric operations to farmers. This motor helps farmers meet the challenge of serving irrigation loads in remote areas. Typically, any irrigation load greater than 15 hp can be served only by a 3-phase motor. Most irrigation sites, however, are several miles from the nearest 3-phase service, and it is seldom economically feasible to extend service to this distance. Solutions to the problem include the use of a phase converter (discussed in Sec. 21.3.8) to create 3-phase power, or installation of gasoline or diesel engines. These solutions can result in poor power quality and can be expensive and time consuming.

A new type of motor has been developed that is slow-starting and can provide up to 60 hp with a single-phase power supply. In addition, this motor can ride through brief power outages. This new type of motor is designed with high-starting torque, and contains magnetic poles which are continuously and instantaneously written on a magnetic layer in the rotor by an exciter pole in the stator. The magnetic poles can be written to a different spot on the rotor during each revolution whenever the rotor speed changes, keeping the pole pattern constant. They are held in the same pattern when the motor reaches full speed. This variation of the poles during start-up gives the motor its slow-starting, high-torque characteristics. A squirrel-cage winding in the motor also adds induction torque in starting.

This slow-starting capability of the motor, combined with the high-starting torque, offers some benefits to the user. The power quality impact on other customers on the line is limited, and the motor pulls only about 2 times the full load current on start-up, compared to a usual draw of 6 times full load current with most motors.

21.3.3 Automation

Larger-scale farming has adopted automation as standard practice. Solid-state electronic devices are used to control livestock feeding by triggering food-release mechanisms at established intervals; phototransistors are used to thin crops by scanning planted rows with precision and at high speed; electronic sensors are also used on farming machinery to monitor shaft speeds, materials flow, and other parameters.

Planting and harvesting machines are equipped with electronic monitoring devices indicating information such as shaft speeds, material flow rates, temperatures, and seeding malfunctions. Reed switches or hinged-plate switches are used to flash light signals or actuate buzzers and horns, and miniature electrical generators, the output of which is proportional to the speed, indicate shaft revolutions per minute relative to a desired value. This information is displayed on a console within the combine cab, or on the tractor.

Electronic crop thinners, using a phototransistor scanning system for each crop row, have resulted in higher vegetable crop yields than is possible with hand thinning, since earlier thinning and a more uniform plant population are possible. Other uses of electronics include automatic temperature and humidity controls for crop drying and storage, and automated surface sprinkler systems such as the solid-set, permanent-overhead, and center-pivot type (utilizing control consoles, operating through buried cables, microwave channels, VHF radio, and pressure and temperature control switches) to provide round-the-clock irrigation control. Overriding time controls provide cooling when predetermined temperatures are reached.

Livestock-feeding systems use electronic controls for time-interval feeding of individual animals based on current production level, weight, age, etc. A transmitter at the feed station sends a signal to a transponder unit on the animal, which upon activation switches on a relay controlling the feed unit. Automatic data recording can be accomplished for individual animals by means of a special neck band, thus enabling rapid and detailed collection of data which, aided by a computer, facilitates efficient management of larger enterprises. Electric motor drives for farm tractors are possible in the near future.

21.3.4 Farm Structures

Water Systems. Water requirements for the farm household and farm enterprises, excluding irrigation, are frequently supplied by a single well. The water-supply equipment is usually an automatic hydropneumatic or air system having pumping capacity of 300 to 600 gal/h and using a $1/4$- to 1-hp motor, depending on the total head in feet and the rate of pumping. Home water requirements average 50 gal/(person)(day). In addition, livestock requirements must be added: each horse, steer, or dry cow, 12 gal; each milk cow, 35 gal for drinking and washing equipment; each hog, 3 gal; each sheep, 2 gal; each 100 chickens, 8 gal. For yard fixtures, each $3/4$-in hose outlet requires 300 gal/h.

Where the source of supply is not more than 22 ft below the pump, a shallow-well system can be used. A jet-centrifugal pump has a practical lift limit of 80 to 100 ft, and piston-type pumps can go as deep as 800 ft with a suction lift below cylinder of 22 ft. This type is placed directly over the well and is generally recommended where pumping depths exceed 80 ft. Automatic pressure switches are usually set to start the pump when the pressure falls to 20 lb and stop it when 40 lb has been obtained. The energy requirement per 1000 gal of water pumped rarely exceeds 2 kWh.

Heating Systems. Electrical heating of farmstead structures is generally confined to milk houses, individual pen-type areas for young livestock, and poultry brooders. Electric heating in the milk house is ideal, as it is odorless, is conveniently controlled, and meets the high sanitary standards required. The milk-house temperature should not exceed 40°F. Several types of heaters have been successfully used: (1) the forced-air circulating type requiring 1500 to 3000 W; (2) batteries of 250-W infrared heat lamps directed toward working areas and water pipes; and (3) heat-pump systems, which utilize the heat removed in cooling the milk. In this type the ice-bank refrigeration system (either bulk or immersion coolers) extracts heat from the water in building up the ice, the heat thus being available for the milk house. Electricity used in this indirect manner produces about three times as much heat as it would if directly used in a resistance heater. Only coolers with 1/2-hp or larger motors are recommended for this application.

In the colder regions, the milk house must be insulated for the most economical cost of installation and operation. In these areas an electrically heated milk house needs at least a 1500-W heater serviced by a 230-V line. Thermostats are usually attached to the heater unit, and operating consumption ranges from 1000 to 3000 kWh a season.

The need for infrared heat lamps during the first week of hog farrowing and sheep lambing has been proved. A 250-W lamp will heat an area 24 in in diameter when 3 ft above the floor. The lamps should be positioned at least 6 in above animals and at least 30 in above the floor when bedding is used.

Ventilation Systems. Electrically powered mechanical ventilation of livestock structures provides low-cost positive control for the removal of excess animal body heat, objectionable odors, and condensation, and for temperature and humidity control. A full-grown cow will give off 3000 Btu/h of body heat; 1000 chickens, about 800 Btu/h. Accurately controlled tests with dairy cows at the University of Missouri showed that temperatures above 75°F and relative humidities over 75% resulted in sharp declines in milk production and body weight.

In general, summer ventilation should maintain inside temperatures equal to or below the outside temperature, while in winter the reverse is true. Thermostatically controlled motor-driven fans are installed as required, with adequate fresh-air intakes to prevent excessive energy costs. Two-speed fans, chosen to move the maximum air volumes required for various livestock, will permit airflow to be reduced in cold weather. Fan motors range from 1/20 to 1/2 hp and will consume 250 kWh/year and up, depending on usage. One kilowatthour of electricity will move about 1 million ft^3 of air.

21.3.5 Plant Production

Irrigation Pumping. More electrical energy is used for irrigation pumping than for any other field operation. Proper design of an irrigating system will depend on the following factors: (1) the acreage and kind of crop to be irrigated; (2) the amount of water that must be supplied; (3) the amount of underground water available; and (4) the depth at which it is found.

Except where the water requirements are small and the depth to water great, plunger pumps are rarely used. The more common type is the centrifugal turbine pump, but where the lift is not more than 15 ft, the horizontal centrifugal pump is also used. The bowl of the turbine pump should be set below any expected drawdown in the well, and this will depend on the porosity of the surrounding strata as well as the rate of pumping.

Vertical turbine pumps require vertical motors with either solid or hollow shafts and thrust bearings capable of carrying the pump load. Horizontal pumps should be connected to their motors through flexible couplings to avoid the use of belts. With average allowance for evaporation, irrigating an acre 1 ft deep requires 340,000 gal. The soil can be wet to a depth of 4 ft by using 4 to 6 in of water. From 10 to 20 in is required to produce the ordinary crops. With an overall efficiency of 50% for pump and motor, each acre-foot of water will require about 2 kWh of electricity for each foot of lift. New motor designs promise to reduce the costs of irrigation for fields remote from the main electrical service.

Methods of Irrigation. These include overhead pipes, stationary spray plants, and portable sprinkler systems. In the overhead type the discharge pipes are supported on posts and are located about 50 ft apart in lengths up to 600 ft. The pipes are usually supported on rollers so that they can be oscillated by a type of water motor, and nozzles are spaced 2 ft or more apart. Sixty gal/min of water per acre at 30 lb pressure is satisfactory. Stationary spray plants can reduce spraying time in orchards by 50% or more compared with portable units. A central pumping station, mixing tanks, and symmetrically located discharge pipes complete the layout. The pumps are usually three- or four-cylinder, single-action, with capacities of 10 to 60 gal/min at pressures up to 600 lb or more, requiring motors of 5 to 30 hp. Outlets are located at regular intervals for attaching the spray hose. Spray nozzles discharge up to 8 gal/min depending on pressure and orifice size. Power required is usually under 10 kWh/(acre)(application). Portable systems utilize lightweight, quick-coupled pipes, with sprinklers attached. Laid on the ground, they require considerable labor to move, but the initial investment is less than with other types. Sprinklers operate at pressures of 20 to 50 lb/in^2 and cover circles 40 to 90 ft in diameter, delivering 3 to 30 gal/min. A motor as small as 2 hp will apply 1 in water to 3 acres of land per week, although larger outfits are commonly used.

Grain Conditioning. Field harvesting and on-the-farm storage losses of small grains and ear corn can be materially reduced where mechanical crop-drying or conditioning equipment is utilized. Early harvest reduces field losses due to shattering or lodging of grain and shelling, which may occur during mechanical harvesting. Crops can be harvested when weather conditions are most favorable as soon as possible after they mature, thus reducing the chance of storm damage while the crop dries in the field.

Heated-Air Crop Dryers. Equipment needed includes an oil burner, a power-driven fan, and a drying bin for the ear corn or small grain.

Most of the dryers are portable. Each unit consists of a power-driven fan, a heater, and safety controls. Such dryers have two characteristics that determine their performance in drying grain: (1) the rate at which heat is supplied (rate of fuel consumption per hour) and (2) the rate of air supply in cubic feet per minute. These dryers are normally equipped with oil burners that consume fuel at the rate of 3 to 14 gal/h and fans powered by 3- to 5-hp electric motors that deliver 9000 to 15,000 ft^3/min of air. Usually 9000 ft^3/min of 30°F air, with a relative humidity of 70%, can be heated to 70°F with an oil consumption of 3 gal/h used in a direct-heat dryer and 4.2 gal/h for the dryer if a heat exchanger is used. The U.S. Department of Agriculture reports that 1000 bu of car corn was dried from 30% to 13% moisture in 167 h.

Shelled corn, wheat, and oats can also be dried with heated air. Depth of grain in drying bins is 4 to 5 ft. Airflow must be uniform through grain, and temperatures of heated air should not exceed 110°F for seed corn and 140°F for wet milling. Temperatures up to 200°F have been used without affecting feed value.

Unheated-Air Crop Dryers. Wheat, oats, and barley are harvested in the summer, when atmospheric conditions are relatively favorable for grain drying with unheated air. Wheat combined at a moisture content as high as 20% can be successfully dried with unheated air. Minimum airflow is 3 ft^3/min · bu with grain up to depths of 4 ft. With 16% moisture content, airflow may be as low as 1 ft^3/min · bu with wheat up to depths of 8 to 10 ft.

21.3.6 Materials Handling

Conveyers and Elevators. Livestock and crop production requires much time and labor for loading, transporting, and unloading materials. Portable chain and flight conveyers, commonly called elevators, are available in lengths of 8 to 50 ft or more and in widths of 6 in to more than 20 in. They may be operated at angles up to 70°, depending on the material being handled, but care must be taken to prevent overturning or collapsing, particularly at the greater angles. The smaller sizes are generally used for moving loose, bulky materials such as small grains, chopped forage, and bedding and will require up to $^3/_4$-hp motors. Larger sizes are mounted on wheels and are used for baled, bagged, and packaged products, as well as other materials. Power requirements range from $^1/_4$ up to 5 hp, depending on the speed, angle of elevation, and weight of the material being handled. Vertical elevators for baled hay are mounted directly to the outside of barn walls. A 42-ft model will require a 2-hp motor.

Auger conveyers requiring fractional-horsepower motors are used for the horizontal and vertical moving of grains. Automatic feeding arrangements may employ 10-in-diameter forage augers in multiples of 5- or 10-ft sections up to 100 ft in length. Three-horsepower motors are required for lengths up to 90 ft, and 5 hp is needed for longer units. *Pneumatic conveyance* of grains and feed is increasingly popular on farms where the distance between storage or processing areas and feeding areas is considerable. This method is safe, has few moving parts, and is dust-free. The pipe can be placed in almost any path, above- or belowground. An air velocity of 4000 ft/min is required for proper operation with a 5-in pipe conveying about 4500 lb of grain/h. This will require $2^3/_4$ hp for each 100 ft of length.

Silo Unloaders. Mechanically operated silo unloaders remove the silage from the silo and deposit it at the foot. The operating mechanism of the top-unloading type is essentially a radial beam with scrapers or augers which collect the silage and bring it to the center of the silo, where it is picked up by a motor-driven air or mechanical device and delivered to the silo chute. Silage then falls down the silo chute, where it is collected for feeding.

There is also a bottom type of silage unloader. The operating mechanism consists of an endless chain mounted on a movable beam. The chain is equipped with scrapers which move the silage out of the silo as the chain revolves.

Unloaders eliminate the need for climbing the silo daily, reduce spoilage by removing silage at a uniform depth, and save up to 200 h/year of time. Results of Ohio State University tests indicate that top removal of grass silage at a rate of 1 ton/h requires 4.3 kWh and that 1.6 tons/h of corn silage requires 2.5 kWh. Three- to ten-hp motors operate the unloaders, and approximately 300 kWh is used annually.

Barn Cleaners. Electrically operated mechanical devices remove manure from poultry, dairy, and livestock barns. In poultry houses the cleaners may be placed under a slatted floor or in a wire-covered pit under tiers of mechanical feeders and waterers. In dairy barns they are installed in the gutters behind the cows. The dragline type uses a motor-driven drum to pull a belt or chain conveyer, equipped with cross flights, to an inclined elevator at the end of the barn, depositing the manure in a field spreader or pit.

The endless-chain type is well adapted to the larger stable where two rows of cows are housed. A single chain with wood or steel paddles travels around the gutters and up a short elevator, discharging the manure outside the stable. In this type of installation connecting or cross gutters must be installed at each end of the two rows of existing gutters so that an endless chain can be installed. The oscillating type uses a reciprocating bar with hinged paddle or auger conveyer. Portable types generally use a scoop steered by the operator and drawn along the gutter by a cable attached to a motor-driven drum. Cleaners are operated by electric motors of 2- to 5-hp capacity. They can be set to operate automatically for a predetermined cleaning period or can be switched on as need arises. Electric-energy use ranges from $^1/_2$ to 1 kWh a month for each cow housed in the stable.

21.3.7 Maintenance

Emergency Power. With increased dependence on electric power for time-controlled mechanical feeding, pipeline milking systems, manure removal, etc., the added investment in emergency power units may be justified compared with the possible economic loss if regular power fails. Generators ranging from 3 to 15 kW and rated at 120/240 V are available in tractor power-takeoff (PTO) and engine-driven types. The latter may be manually or automatically started. Automatic generators must be of higher capacity, because peak-connected loads will be carried if power fails. Nonautomatic types should have a "power off" alarm and must be PTO-equipped with an overload circuit breaker. The tractor PTO-driven generators are least expensive to purchase, as the tractor engine serves as the generator drive. Output is controlled by an engine tachometer and/or voltmeter in the generator unit. Manufacturers claim a voltage rating within 2% of the normal supply voltage. Required generator capacity is obtained by totaling the power needs of essential loads, plus allowances for future loads and high starting currents of the motors. Double-throw switches must be used at the point of connection into the wiring, to prevent generator damage and power feedback into the supply line.

Arc Welders. A highly mechanized agriculture requires that many machinery and structural repairs be made by the farmers themselves. A survey by the Kansas Farm Electrification Council indicates that the number of dollars invested in electric welders was greater than for any other item of electric farm equipment. The electric arc welder is inexpensive, efficient, and an almost indispensable tool on modern farms. The 180-A transformer-type ac welder is satisfactory for most farm shops. This machine can cut, hard-surface, and weld metals up to 1/2 in thick. It requires a line voltage of 220 to 240 V single-phase, 60 Hz. Current outputs from 30 to 180 A are possible. Duty cycle at maximum output is 20% with an open-circuit voltage of 25 V. A carbon-arc-torch attachment is used for brazing, soldering, and heating purposes. Larger generator-type units, either engine or tractor PTO-driven (hence portable), may be used as emergency power generators, supplying 5000 W of 230- or 115-V single-phase 60-Hz power.

Phase Converters. Most farms have 100- or 200-A single-phase service, which limits them to the use of 7 1/2- or 10-hp motors. Two types of phase converters are available which will convert single-phase to 3-phase current. By connecting the converters between the electric meter and the motor, they will permit the use of 3-phase motors up to 20 hp or more. In addition, the *National Electric Code* states that service entrances need be heavy enough to handle only the largest-power-demand equipment, plus a portion of all other equipment, rather than the total connected load as before. This is advantageous where irrigation pumps, grain dryers, large feed mills, etc., are in use.

Other Shop Equipment. This includes electrically powered air compressors, drill presses, grinders, hoists, lathes, saws, and paint sprayers. These generally require 1/4- to 1/2-hp motors. Battery chargers drawing approximately 2 kWh/charge are popular.

21.4 THE FOOD INDUSTRY

The food industry is diverse, essentially composed of eight sectors: dairy, processed fruits and vegetables, breakfast cereals, wet corn milling, bakery products, sugar and confectionery, fats and oils, and alcoholic beverages. The industry relies on a variety of energy sources.

Food manufacturing processes are as diverse as the industry itself. They can be as simple as boiling a potato or as complex, intricate, and technical as converting corn into starches, sweeteners, and corn oil. In between these two extremes are numerous processes with varying degrees of technological complexity, requiring a mixture of physical, thermal, and biochemical transformations.

One of the simpler food production processes is liquid heating, essentially applying heat through water or some other medium. For example, water temperatures of 125°F and above are required for the scalding of poultry, while pressure cooking involves the application of steam in rotation cookers. In the beverage industry, boiling converts insoluble starch into liquefied starch to obtain brewer's malt. Baking involves the application of heat to cook or brown foods. Baking and roasting apply heat directly to the foods. Frying uses a transfer medium—typically oil. Baking is used extensively in the bread-and-cookie industry. The more complex processes also involve the application of thermal energy to effect chemical or biochemical changes in the composition of foods. Temperatures above 140°F are required to pasteurize milk.

Food preservation techniques include thermal processing, moisture removal, chemical processing, and irradiation. Cooking, boiling, baking, and other procedures use heat to inactivate microorganisms, enzymes, or harmful chemicals in food. Removal of some of the moisture in foods often extends the length of time during which the product can be stored without spoilage, such as in dehydration. Current dehydration techniques include warm, hot air, or steam; freeze drying for coffee and herbs, and the concentration of milk and fruit juices; and microwave and vacuum technologies. The chemical additives most commonly used with food are salt and smoke (used in curing and smoking operations). The last type of food preservation technique is a recent introduction: ionizing radiation—or irradiation—which uses gamma rays, electron beams, and x–rays to sterilize a variety of products. This sterilization is brought about by a chemically induced destruction of insects and microorganisms which would otherwise accelerate the spoilage process. The industry is critically dependent on energy for food preservation.

The most electricity-intensive food-processing sectors are wet corn milling, meat-packing plants, fluid-milk producers, soybean-oil mills, malt beverages, frozen fruits and vegetables, poultry dressing plants, bakeries, flour-milling plants, and soft-drink plants. Most of these consume large amounts of electricity because of their size.

21.5 THE TEXTILE INDUSTRY

The textile industry consumes raw materials from the agricultural and chemical industries and supplies raw materials to producers in the home furnishings (accounting for 45% of industry output), apparel (30%), and industrial fabrics industries (25%). Overall energy consumption in the textile industry has increased in response to industry growth and productivity improvements. Electric machine drives account for 57% of electricity use, lighting and other facility services consume 23%, and HVAC systems account for about 14%.

In response to increased foreign competition, the textile industry has had to invest in productivity and efficiency improvements. Productivity improvements are reflected in increase in the amount of energy consumed per employee. This increase is largely attributable to increased process automation. Improvements in energy efficiency are evident in the decreased amount of energy consumed per value of shipment. The textile industry uses large amounts of energy in process heating. Although large electric and thermal loads provide attractive opportunities for cogeneration, cost pressures have constrained cogeneration investments.

In response to highly competitive market conditions, textile manufacturers have adopted “quick response” (QR) systems that improve communication between suppliers, manufacturers, and end users. Although QR systems are fundamentally information technology systems that are not themselves energy intensive, they provide opportunities to increase the level of automation in a facility.

The most electricity-intensive textile plants are greige mills, which produce unfinished woven and knitted goods. Electricity supplies about 80% of the total energy requirement in a greige mill. The application of mechanical moisture-removal equipment such as vacuum extractors and roller squeezers is expected to impact electrical plant design and consumption, primarily as a result of the potential energy savings and process simplification in textile drying and finishing.

21.6 THE PETROLEUM INDUSTRY

The petroleum refining industry encompasses a broad range of operations that provide raw materials to almost every manufacturing industry. Petroleum refiners produce fuels and solvents—lubricants that are essential to the operation of most industrial facilities. Additionally, petroleum refiners produce feedstocks for the plastics and petrochemical industries.

Electricity accounts for about 3.5% of the industry’s total energy use. About 80% of the total electricity consumption is used to operate pumps, compressors, fans, and other machine-driven applications.

21.6.1 Oil Refineries

Introduction. Oil refineries are seeking to maintain or increase their profit levels by upgrading existing processes and replacing aging equipment with more efficient and lower-operating-cost options. In the petroleum refining process, the primary raw material, crude oil, is heated, separated, and converted into up to 10 major product categories: motor gasoline, fuel gas, liquefied petroleum gases (LPGs), jet fuel, diesel fuel, lubricants, kerosene, fuel oils, asphalts, and coke.

In the United States, motor gasoline and jet fuel provide the highest revenue and are produced in the highest quantities. Diesel and home heating oils and residual fuels used by power plants, ships, and industrial boilers are also important products that contribute to the industry's bottom line. Each refinery has a varied product slate, the mix of products produced in response to market demands.

Generally speaking, a refinery can be classified as one of three types: gasoline (cracking), fuel oil (hydroskimming or topping), and coking. A gasoline refinery produces a greater percentage of gasoline and less heating and other heavy oils. The refinery increases its gasoline yield by upgrading the heavy oils. A fuel oil refinery primarily produces ship, industrial, and home heating oils, some gasoline, and kerosene. A maximum amount of gasoline and a much reduced percentage of heavy oils are produced in a coking refinery.

General. Oil refineries vary greatly in the variety and quantity of products and crude-oil throughput. A basic refinery producing gasoline and other fuel products would include operations such as those described below. These are merely typical of the many processes in current use.

Crude-Oil Desalting. Water is added to the crude oil to dissolve the unwanted salt. The mixture is passed through a vessel containing electrodes between which a potential of several kilovolts is maintained. The potential gradient causes the salt and water to coalesce and settle to the bottom of the vessel, where the mixture is drawn off. The desalted crude oil is discharged near the top of the vessel.

Crude-Oil Distillation. The crude oil, which is a mixture of a large variety of hydrocarbons having different boiling points, is heated in a furnace to about 750°F and then enters a fractionating tower. The components are separated according to boiling range, since the lighter ones rise in the tower as gases and the heavier ones fall in the tower as liquids. Trays with specially designed openings in them are installed at intervals in the tower to ensure intimate mixing of the rising gases and the falling liquids and to provide places where liquids having certain boiling ranges may be drawn off the tower. The operation is first performed in a distillation tower in which the pressure is maintained somewhat above atmospheric pressure and again in another tower which is kept under vacuum in order to reduce the boiling temperatures of the hydrocarbons and thereby prevent thermal cracking, which is decomposition due to excessive heat. The combined unit is called an atmospheric and vacuum, or A&V, unit. It is also sometimes called a "two-stage pipe still." The main fractions produced are condensable gas, gasoline components, diesel fuel, heating oils, gas oils, and residuals. All of these are processed further.

Fluid Catalytic Cracking. To increase gasoline yield from a crude oil, heavy gas oil from the distillation unit is processed in a cracking unit. The heavy molecules are brought in contact with a catalyst under proper conditions of temperature and pressure and are converted into lighter molecules. Thus lighter products are formed which are suitable for use as gasoline and distillate fuel components. The use of a catalyst promotes the cracking reaction at a lower temperature and pressure and produces larger quantities of products having more valuable qualities than is possible with straight thermal cracking.

The clay catalyst is in powder form, and it is handled as a fluid. The cracking reaction causes the formation of carbon deposits on the catalyst particles. These are removed by controlled burning in a regenerator vessel. The catalyst is continually being circulated through the reactor and the regenerator by means of gas flow and airflow, respectively.

Combustion Air Blower. Combustion air for the regenerator is provided by a large centrifugal air blower driven by an induction motor and gear increaser or by a steam or gas turbine.

Gas Compressor. Some of the products of the catalytic cracker are drawn off as gas. This gas is compressed, condensed, and fractionated to provide other fuel products and feedstocks for petrochemical processes. A centrifugal compressor is used, and it is driven by an induction motor through a gear increaser or by a directly coupled steam turbine or gas turbine.

Catalytic Re-forming. In this process hydrocarbon molecules are rearranged and recombined to form molecules of higher octane rating which can be used as gasoline components. Hydrogen is produced in this process. Large reciprocating and centrifugal compressors are used to move the large volumes of gas involved in the process. Hydrogen is a by-product of this process.

Hydrofining. This process uses hydrogen in the treatment of other products, such as jet fuels, distillates, and lubricating oils, to improve quality and to remove sulfur.

Cooling-Water Pumps. Large volumes of water are used for cooling-process streams and for condensers. Water may be conserved by the use of induced-draft cooling towers.

Cooling-water circulating pumps are driven by vertical motors. Standby pumps are driven by steam turbines.

Power Supply. Refining operations are continuous processes, and uninterrupted runs of 1 or more years are expected between planned shutdowns for maintenance or turnarounds. Therefore, it is essential that the power supply be extremely reliable.

Often duplicate full-capacity feeders are installed to the refinery, and sometimes these are run from different substations, for increased security. Many refineries practice cogeneration to take advantage of their high steam loads.

Distribution Systems. Because refinery loads are often concentrated to a large extent in fairly well-defined process areas, it is common to install unit substations in the major process areas. These substations contain power transformers to provide 4160- or 2400-V power, 480-V power and lighting transformers, or provision for feeding lighting transformers. They also contain all associated switchgear, motor-control, and emergency generators.

A typical refinery *process-unit distribution* system is shown in Fig. 21-2.

Area Classification. Flammable gases and vapors are processed in oil refineries. Therefore, it is necessary to classify the various locations according to the material that is present and also according to the degree of hazard expected.

Reference should be made to the applicable electrical code for requirements governing installations in classified areas.

The actual classification of areas is usually made by the electrical design engineer in consultation with persons who are familiar with the operation of the process.

The most common classification for refinery process units is Class I, Group D; the class consists of hazardous gases and vapors, and the group comprises gasoline and many of the petroleum products.

Current practice is to classify outdoor, freely ventilated process areas as Division 2. Indoor process areas that are not freely ventilated and places below grade level are classified as Division 1. Areas in which a permanent ignition source is located, such as around a furnace, are not classified. Pressure-ventilated unit substations and control buildings are not classified. However, some companies follow the practice of classifying control rooms as Division 2.

21.6.2 Electric Motors

Type of Motor. Two-pole induction motors having NEMA Design B characteristics are used to drive the majority of refinery-process pumps. Motors operating at slower speeds are used for some applications, such as for driving reciprocating compressors or for driving centrifugal compressors through gear-speed increasers.

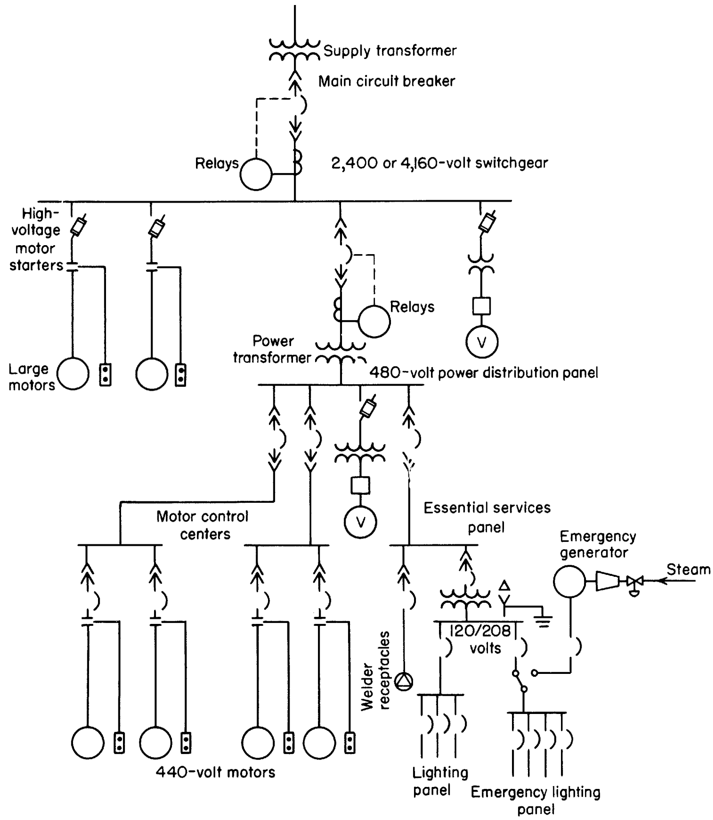


FIGURE 21-2 Typical refinery process-unit distribution system.

21.6.3 Emergency Power Supply

Emergency Generator. Generators provide emergency power for critical motor loads, emergency lighting, and instruments. Generators are driven by steam turbines or diesel engines, which start automatically when the normal refinery power supply fails. Capacities range from 15 to 300 kW. Output voltage is 120/208 V 3-phase or 480 V 3-phase. Load is automatically transferred to the generator after it has reached normal operating speed.

Batteries. Batteries are used when a continuous supply of power must be available for process instruments, emergency controls, and shutdown devices. They are also used for remote-control systems. Voltages of 24, 48, and 120 V are used. Inverters provide power for critical ac instruments and are sometimes used for computer power supply. Battery chargers are transferred to the emergency generator on loss of normal power.

Batteries should have enough capacity to carry full load for half an hour with the charger off. Charger capacity must be sufficient to carry the full dc load and to recharge the battery in 8 h.

21.6.4 Oil-Well Pumping

Methods of Forced Production. When there is insufficient natural pressure to force the crude oil to the surface, some method of forced production is used. The most common methods are described as follows.

High-Pressure Gas Lift. Gas is forced to the bottom of the well or to an intermediate point in the well. The gas mixes with the oil in the well and induces flow by decreasing the density of the fluid.

Water Flooding. Treated water is forced into the formation through nearby wells in order to increase the pressure in the formation and induce flow.

Bottom-Hole Hydraulic Pump. High-pressure crude oil is carried down the well in tubing, and it is used to drive a reciprocating pump located at the bottom of the well.

Bottom-Hole Centrifugal Pump. A special motor-driven multistage centrifugal pump is lowered to the bottom of the well. This method is used where large volumes of fluid must be pumped.

Sucker-Rod Pump. A reciprocating single-acting pump is installed at the bottom of the well on the end of a tube inside the well casing. In sucker-rod pump drives, the plunger is operated by a sucker rod from the surface. Various methods are available to drive the sucker-rod string, but generally a walking beam is used to provide the desired vertical motion.

Central Power Units. Central power units driven by electric motors are sometimes used to serve as many as 15 or 20 wells. Operating rods lead out to each pump to provide reciprocating motion to the walking beams.

Individual Engine or Motor Drives. These are more commonly used than multiple drives because they can be started and stopped individually and the pumps can be operated at different speeds. Electric motor drives are preferred because they can be started and stopped by a timer, they provide consistent, trouble-free performance regardless of weather conditions, and maintenance and investment costs are low. Also, it is easy to measure power demand and energy consumption of an electric motor. The well may be counterbalanced readily by an ammeter.

Motor Types. Torque requirements vary widely during the pumping cycle, and peaks occur when the sucker-rod string and fluid load are lifted and when the counterweight is lifted. NEMA Design D motors, although relatively expensive, are well suited to this service, since they minimize current peaks and provide adequate torque under all service conditions, including automatic operation by time control. NEMA Design C motors may be used where operating conditions are less severe. NEMA Design B motors must be used with care in this service to avoid high cyclic current peaks, which may be objectionable on a small system, particularly if several wells should "get in step." The use of Design B motors can also lead to oversizing of motors in an attempt to obtain sufficient starting torque. This results in the operation of the motor at a relatively low load factor, with consequent low power factor.

Double- or Triple-Rated Motors. These are special motors developed for oil-well pumping. They are totally enclosed, fan-cooled NEMA Design D motors that can be reconnected for 2- or 3-hp ratings at a common speed of 1200 r/min. Typical horsepower ratings are 20/15/10 and 50/40/30. They provide flexibility in the field since they permit the selection of the horsepower rating at which the motor may be operated most efficiently. They also permit changing the pumping speed by changing the motor pulley and reconnecting the motor.

Single-Phase Operation. If single-phase power only is available, it is advisable to consider the use of single-phase/3-phase converters and 3-phase motors. This avoids the use of large single-phase capacitor start motors, which are relatively expensive and contain a starting switch which could be a source of trouble due to failure or to the presence of flammable gas in the vicinity of the well.

Oil-Well Control. A packaged control unit is available to control individual oil-well pumps. It contains, in a weatherproof enclosure, a combination magnetic starter, a time switch that can start and stop the motor according to a predetermined program, a timing relay that delays the start of the motor following a power failure, and lightning arresters. Push-button control is also provided.

Power-Factor Correction. The induction motors used for oil-well pumping have high starting torques with relatively low power factors. Also, the average load on these motors is fairly low.

Therefore, it is advisable to consider the installation of capacitors to avoid paying the penalty imposed by most power companies for low power factor. They will be installed at the individual motors and switched with them, if voltage drop in the distribution system is to be corrected as well as power factor. Otherwise they may be installed in larger banks at the distribution center, if it is more economical to do so.

21.6.5 Gas-Processing Plants

Natural Gas. Natural gas varies widely in composition and contains undesirable materials such as water and sulfur compounds, which must be removed before the gas enters the transmission pipeline. Various chemical processes are used, and plant capacity ranges from 5 to 1000 million ft^3 of gas processed/day. By-products such as propane, butane, pentanes, and elemental sulfur are produced and marketed.

Power Supply. Purchased power is used where available. Local generation is by reciprocating gas engines or gas turbines.

Electrical installation practice in gas plants is similar to that followed in oil refineries, as described below.

21.6.6 Oil Pipelines

Gathering Systems. These collect crude oil from the individual wells, or from tanks located near them, and carry it to tankage, where shipments may be accumulated.

Trunk Lines. These feed crude oil from gathering systems to the main crude pipeline pumping station.

Crude Lines. Crude lines are generally operated as common carriers. Since they handle crude oil for several companies and because crude-oil shipments vary greatly in quality and composition, storage tanks are necessary at stations along the pipeline so that batch shipments may be handled.

Product Lines. These convey products from a refinery to the market area. Some products lines are operated as common carriers, while others are privately owned and operated.

Operation of Pipelines. Batches of crude oil or products are dispatched through a pipeline and are withdrawn to tankage at the end of the line or at intermediate points. If a batch is being drawn off at an intermediate point, the downstream stations will operate at reduced flow.

Little mixing occurs at the interface between different batches. By careful scheduling and a knowledge of the pipeline it is possible to predict fairly accurately when an interface will arrive at a station. Interface detectors are installed also.

Pumping Stations. A schematic diagram of a typical pumping station is shown in Fig. 21-3.

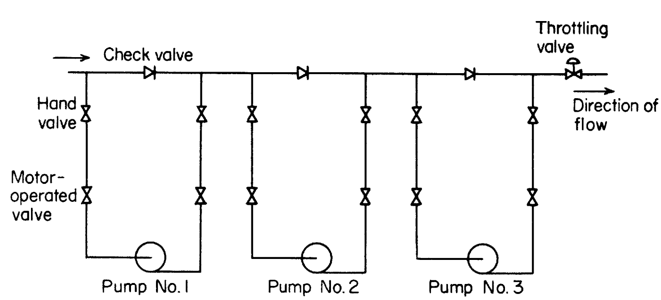


FIGURE 21-3 Pipeline pumping station.

Arrangement of Pumping Stations. These are located at the head of the pipeline and at intervals along the line. Intermediate or booster stations must be capable of operating under varying conditions due to differences in liquid gravity, withdrawals at intermediate points, and the shutting down of other booster stations. Pumping stations often contain two or three pumps connected in series, with bypass arrangements using check valves across each pump. The pumps may all be of the same capacity, or one of them may be half size. By operating the pumps singly or together, a range of pumping capacities can be achieved.

Throttling of pump discharge may also be used to provide finer control and to permit operation when pump suction pressure may be inadequate for full flow operation.

Control of Pumping Station. Pumping stations are often unattended and may be remotely controlled by radio or telephone circuits.

Electrical System. Figure 21-4 is a typical electrical single-line diagram for a pumping station.

Motor Type for Main Pumps. The main pumps are driven by 3600-r/min induction motors having NEMA Design B characteristics. Full-voltage starting is used.

Motor Enclosure. Motor enclosures for outdoor use are NEMA weather-protected Type II, totally enclosed, fan-cooled, or dripproof with weather protection. Motors of the latter type are widely used. Not only are they less expensive than the other types, but they also have a service factor of 1.15. The above enclosure types are all suitable for the Class I, Group D, Division 2 classifications usually encountered.

If the pumps are located indoors, a Division 1 classification is likely to apply. Motors must be Class I, Group D, explosionproof, or they may be separately ventilated with clean outside air brought to the motor by fans. Auxiliary devices such as alarm contacts on the motor must be suitable for the area classification. The installed costs, overall efficiencies, and service factors associated with the enclosures that are available will influence the selection.

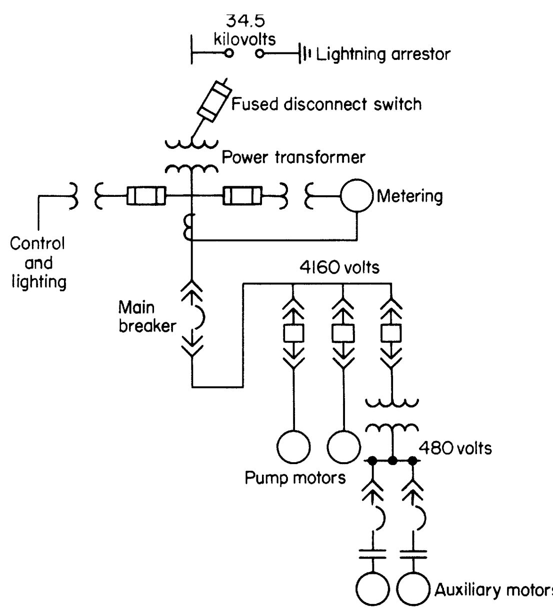


FIGURE 21-4 Single-line diagram of a pipeline pumping station.

Adjustable-Speed Drives. Of the total electricity consumed by petroleum refiners, 80% is for machine-drive applications. Most of this energy is used to drive turbomachinery such as pumps, compressors, and fans. These applications are well suited for adjustable-speed drives. Since the power consumed by turbomachinery is highly dependent on operating speed, the use of adjustable-speed drives to control the output of these machines can provide significant improvements in operating costs and system performance. The most common type of adjustable-speed drive is the *variable-frequency drive* (VFD), which uses rectifiers and inverters to control the frequency of the power supplied to a motor. VFD technology is widely used in industry, providing energy savings up to 70% below conventional flow control options.

Pumping systems are particularly attractive applications for VFDs. Because of the reduced amount of wear on system valves and piping supports that results from minimizing unnecessary pump output, VFDs can improve system reliability and lower maintenance costs.

21.7 THE STEEL INDUSTRY

Industry Description. The domestic steel industry includes blast-furnace integrated steelmakers, nonintegrated minimills, and independent producers of wire, bar, and pipe made from raw steel. Blast-furnace operations use either an open hearth or a basic oxygen furnace, with electric arc furnace used in the minimill. Raw steel is melted increasingly by electric furnaces, and the basic oxygen furnace is still used in a number of mills.

Energy Use. The steel industry is highly energy-intensive. Its aggregated average energy consumption of approximately 19 million Btu per ton of steel shipped represents about 3% of the energy consumed in the United States and 10% of that used by the industrial sector. Approximately 60% of the energy consumed by the steel industry is derived directly from coal, and much of the electricity the industry uses is generated at coal-fired power plants.

Natural gas accounts for nearly 30% of the industry's energy consumption. Energy purchases represent about 15% to 20% of the total manufacturing cost of steel. In response to these significant costs, the industry has sought to improve production efficiency, principally through process modifications, new technology, and the retirement of old or inefficient plants. Since 1975, the steel industry has reduced energy consumption per ton of steel shipped by about 45%. The steel industry blast furnace process lends itself naturally to cogeneration by virtue of the large amount of by-product gases generated. The industry today recycles a high percentage of scrap steel, and uses fuels that are by-products from the cokemaking and ironmaking processes to cogenerate electricity and steam. Additionally, the industry has replaced a number of their open-hearth furnaces with basic oxygen furnaces and has implemented continuous casting on a wide scale.

Electricity as a Cost of Making Steel. The U.S. steel industry is a major consumer of electricity. Energy costs account for 15% to 20% of the total manufacturing cost of producing steel. Electricity on average represents only about 7% of the total energy consumed by the industry, but at some steel plants over half of the purchased energy is in the form of electricity. Costs of electrical energy are disproportionately higher than those for other forms of energy.

Power-Distribution System. Integrated steel mills having blast furnaces and coke plants make use of the combustible gases from these processes by burning them to produce power and process steam. Many older steel plants produce and use power at 25 Hz, utilizing primary distribution voltages of 6.9 to 13.8 kV and secondary systems of 4160 or 2400 V and 480 V. Modern steel plants and modern parts of older plants utilize 60-Hz power exclusively, with primary distribution at 69 or 138 kV and secondary voltages of 13.8 kV, 4160 or 2400 V, and 480 V. Power can be supplied by public utilities, by in-plant generation, or by a combination of the two. Some plants having both 25- and 60-Hz systems have conversion equipment, typically large Scherbius sets or rectifier-inverter systems, for transfer of power from one frequency to the other. Where power is supplied from both a public utility and in-plant generation, there is usually some provision for controlling the maximum demand and improving the power factor of the portion of the load supplied by the utility to avoid high penalty charges.

The presence of numerous cranes and other equipment requiring dc motors with some control over speed has led to the extensive use of 250-V constant-potential dc shop circuits in most steel mills. In older plants the direct current is supplied by rotary converters, motor-generator sets, or mercury-arc rectifiers. In modern plants it is supplied by silicon diode or thyristor rectifiers. The trend is toward the elimination of dc shop circuits by the use of ac cranes and package thyristor power supplies for drives requiring variable speed.

Primary Production. The basic steelmaking areas of an integrated steel plant consist of *coke-oven batteries* for conversion of coal to coke, *blast furnaces* for conversion of iron ore to molten iron, and *steel-producing units* for refining molten iron and other alloy ingredients to steel. Once this basic steel has been produced in ingot (block) form or "continuous cast" into semifinished bars, it is ready for subsequent rolling into a usable size and shape.

Power consumption per ton of steel produced is low in the basic steelmaking areas because the products arc handled in molten or bulk form, compared with the rolling mills, where reheated or cold steel is literally squeezed and stretched to the desired size and shape. Much of the electric power consumption in these primary producing areas is associated with auxiliary drives involved in material handling, water, air, and by-product utilization, and mobile equipment.

The processes of iron reduction and steel refining, with their many, sometimes elusive, variables, do lend themselves to automatic and computer control. Open-loop computer systems have been applied to *blast-furnace* and *basic-oxygen-furnace* (BOF) operations. Raw-material handling and charging functions in *blast-furnace stock houses* and BOF have been automated extensively.

In a minimill, which includes specialty shops, the basic raw material is steel scrap which is melted and refined to produce raw steel and steel products. The steel is melted by an electric arc in which the current passes from one electrode through an arc to the scrap charge, and then from the charge to another electrode. The molten steel is then refined by rapid oxidation and the reaction of impurities with added slag materials.

Rolling Mills. Rolling mills are classified either according to their construction or according to the material processed. Classified according to construction, mills are generally two-high or four-high, with a few existing three-high mills. Four-high mills consist of the usual two work rolls in contact with the product, with two additional "backup" rolls which are much larger and allow high rolling pressure without excessive deflection of the work rolls. A "universal" mill has vertical or edging rolls in tandem with the horizontal rolls. This permits a reduction of width or control of the edges of the product in the same stand where a reduction in thickness is taking place.

The principal types of mills, classified as to product rolled, are as follows.

Blooming Mills. These mills roll ingots into blooms, or slabs. All material rolled in steel mills, except that which is direct continuous cast into slabs, blooms, or bars, first passes through this type of mill, or its equivalent, to be reduced to proper dimensions for handling in the finishing mills. These are generally single-stand, two-high reversing. Slabbing mills are a modification of the blooming mill. They are usually universal mills, which permits convenient rolling of wide slabs by eliminating frequent turning of the ingot. Blooming and slabbing mills may have such automatic features as preset of roll openings and speed synchronization between main and edger rolls. The preset information is sometimes stored on business-machine cards and read into the mill control system by a card reader, or it can be stored in the computer memory in computer-controlled installations. Blooming and slabbing mills are powered by large low-speed dc motors operating from a variable-voltage system. A typical slabbing mill has a total of four 3000-hp motors driving the horizontal rolls and two 2000-hp motors driving the edger rolls at rated motor speeds of 40 to 80 r/min. The dc power for blooming mills has been traditionally supplied by generators using the Ilgner system, but the present trend is toward the use of thyristor power supplies connected in rectifier-inverter configurations.

Hot-Strip Mills. These mills roll sheets, strip, and plate from heated slabs. These mills can be placed in two categories: continuous and semicontinuous. (These terms designate the type of rougher which precedes the finishing train.) The *continuous mill* has two to six mills in line which reduce the slab to a predetermined thickness for subsequent rolling through the finishing train.

The *semicontinuous mill* has a reversing rougher on which reduction is made by running the piece back and forth through the mill, which reduces the slab to a predetermined thickness for subsequent rolling through the finishing train.

The finishing train consists of five to seven stands, closely coupled and synchronized in speed, in which the piece is reduced to the desired gage. As a general rule, the piece is in all finishing stands simultaneously.

Roughing stands of a continuous mill are usually driven by ac motors, since speed synchronization with an adjacent stand is not required. Many existing roughing mills utilize wound-rotor induction motors with flywheels and water-slip regulators, but today synchronous motors predominate. The roughers on semicontinuous mills are driven by dc motors.

Direct-current motors operating from a variable-voltage system have been traditionally used for finishing stands, but ac motors with variable-speed drives are now being used in some applications. Tension between stands must be accurately controlled at a relatively low value, since the hot steel is in plastic form and excessive tension results in "necking," or breaking. Constant-tension "loopers" are used between stands for interstand tension regulation. The looper is pushed up by air, hydraulic pressure, or a torque motor, and the upward movement is restricted by the strip. The looper position is then fed back to the electrical control system, and the speed of individual finishing-stand motors is regulated to maintain the proper looper position.

Various measuring devices must always be in operation, measuring gage, screw position (roll opening), looper position, roll force, and width of strip. With the feedback from these devices, the positioning of the screw-downs, which control gage, and the edgers, which control width, can be regulated during the rolling of the strip. The newer mills make use of digital positioning to set up the mills prior to the entrance of the bar into the mill.

The thread-speed reference is set into the speed controller, and this speed is maintained by utilizing the pilot-generator feedback. When the strip enters the mill, the load increases sharply, bringing the current feedback to the armature controller into effect. By controlling the armature and field, rapid speed changes while complete control is maintained are possible, both with and without load. The acceleration control, when called on, feeds into all the finishing-mill motors for uniform acceleration.

Newer mills now have small digital computers controlling each stand.

As the strip cools during rolling, it becomes harder and the output gage changes. Part of this change in gage can be controlled by accelerating the mill, but much of it must be controlled by changing the position of the screw-downs. The newer hot-strip mills use automatic gage control systems (AGC) to regulate screw-down position while the strip is in the finishing mill. There are two types of AGC in use, a constant-gage system and an absolute system. Both systems work basically the same, by taking a reference either from the head end of the strip in the constant-gage system or from preset switches in the absolute system and maintaining it throughout the length of the strip. This reference is compared with the actual gage of the strip, measured as a function of roll force and screw-down position, or by an x-ray gage, and the screw-downs are moved to bring both these signals to the same value. This is a continuous operation during the period of time the strip is in the finishing mills.

Most modern hot-strip mills are built with computer control. Rolling in a completely automatic or computer-controlled state, the computer provides the position and speed references instead of operators. With a computer-controlled mill, these references can be constantly changed as the computer receives new information from the strip.

The handling of steel between roughing stands and between the finishing stands and the downcoilers is accomplished by roll "tables," which are a series of motor-driven rolls on which the steel lies. In some installations, one large table motor drives a number of table rolls through mechanical gearing and shafting ("line-shaft drive"); in others, each roll is driven by one small motor. Permanent-magnet field dc motors have been particularly successful for this arrangement. Table rolls have been driven by dc motors operating from a variable-voltage system, but newer applications use variable-frequency induction-motor system. As adjustable-frequency power supplies have become more reliable, this system is becoming standard.

Tandem Cold-Strip Mills. These are used to cold-reduce previously rolled hot-strip mill products down to thicknesses as low as 0.002 in. Special "foil" mills have been built which can roll even lighter gages. A cold-strip mill is similar to the finishing stands of a hot-strip mill, except that tension between stands plays a much more important role in reducing the thickness of the steel. Modern cold-strip mills are built for finishing speeds in excess of 5000 ft/min.

Cold-strip mills are generally three- to six-stand mills, four-high, with a coil box or payoff reel on the entry end and a tension reel at the delivery end. Newer cold-strip mills are universally driven by variable speed ac motors. Usually the individual stands are voltage-regulated, and the operator establishes the motor field and, thereby, the stand speed according to the gage reduction and the strip tension desired. Load-cell tensiometers are used to indicate to the operator the interstand tensions.

Some of the most modern mills have each stand speed-regulated, and all major drives incorporate full field acceleration to base motor speed for maximum torque. Above base motor speed, automatic field weakening with constant rated armature voltage is used to attain top motor speed. Payoff-reel tensions, interstand tensions, and winding-reel tensions are accurately controlled by using load-cell tensiometers to measure tension and provide feedback to tension regulators which operate on the appropriate drives (see Fig. 21-5).

Finished steel gage is controlled by utilizing x-ray gages to provide information to AGC regulators. Gage control may be accomplished by operating on interstand tensions only or by a combination of interstand tensions and work-roll openings. Work-roll openings ("screw-down positions") are controlled by digital position regulators which hold the screw-down position constant to within 0.0001 in. To help further in producing constant finished gage, the rolling forces on each stand are held constant by using speed-programmed digital regulators to hold constant rolling force at all speeds. Rolling force is measured and indicated by load cells placed under each stand housing.

Other Types of Rolling Mill. *Billet mills,* used to roll blooms into billets, are frequently of a continuous type with rolling stands in tandem with several sets of passes in the rolls so that different-sized billets can be produced from a given bloom without changing rolls. This type of mill is used for producing only a limited range of sizes, which are further reduced in finishing mills.

Plate mills produce plates from slabs previously rolled by a blooming or slabbing mill. These mills are generally single-stand mills, either two- or four-high reversing or three-high running

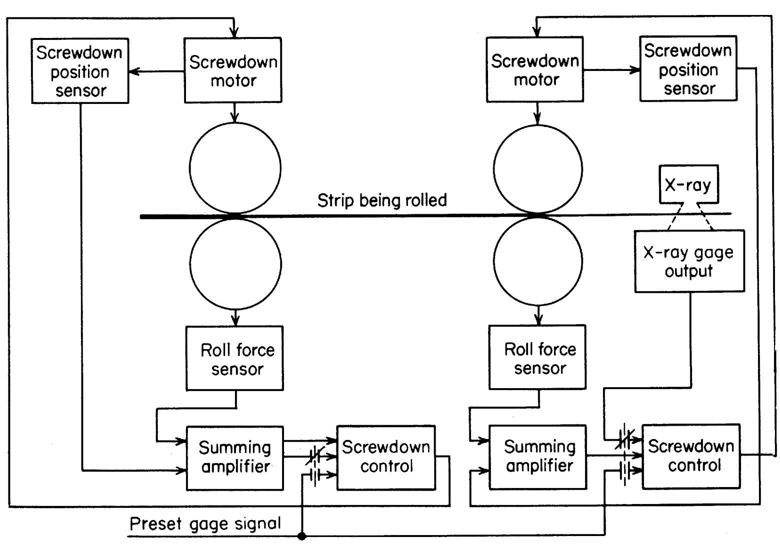


FIGURE 21-5 Cold-strip mill.

continuously in one direction, although some mills are provided with several stands in tandem, one serving as a rougher or breakdown stand. They are sometimes universal mills, provided with vertical rolls for finishing the edges of the plates.

Structural mills are used for rolling beams, heavy angles, channels, etc., from blooms or billets. Such mills rolling the smaller structural sections and miscellaneous shapes are sometimes called "bar mills." They are frequently three-high, with more than one stand in line, and frequently have a separate rougher or breakdown stand. Some also have two-high reversing finishing mills with edger and vertical, as well as horizontal, rolls, to produce wide-flange (H) beams.

Rail mills are special mills for rolling this product only from blooms or billets, although rails are sometimes rolled on structural mills.

Merchant mills are used for rolling small angles, channels, rounds, squares, etc., from billets. This classification is generally applied to mills rolling the smaller sections used for miscellaneous purposes. Most are two-high, in tandem; however, some are arranged for cross-country rolling.

Rod mills are a specialized type of merchant mill for rolling small rounds, usually from No. 5 BWG and upward, which are later drawn into wire. Modern rod mills are arranged to roll a multiplicity of strands simultaneously.

Temper mills are used to produce steel strip of the desired temper, flatness, surface, and luster by using rolling pressure and tension. Reduction in thickness is incidental in the process and is normally very slight. Temper mills generally have one or two 4-high stands and are similar to cold-strip mills, except that they are of lighter construction, have less powerful motors, and have simpler electrical control systems. Temper mills have been built and operated at speeds in excess of 7000 ft/min.

Tube mills are used to produce tubes, for example, pipe and conduit, by either the seamless or the butt- or lap-weld process. Seamless mills pierce a solid billet and then form the pierced billet into tubes of the desired size and thickness. Butt- or lap-weld mills form and weld previously prepared pipe skelp into tubes.

21.8 THE CHEMICAL INDUSTRY

The chemical industry encompasses a broad range of manufacturers that provide essential raw materials for many other manufacturing processes. The largest energy users within the chemical industry are chloralkali facilities, industrial gases, inorganic chemicals, plastics and resins, synthetic rubber, organic fibers, cyclic crudes, nitrogenous fertilizers, and industrial organic chemicals.

21.8.1 Industrial Gases

The industrial gas facilities separate nitrogen, oxygen, argon, and carbon dioxide from air. These gases are then shipped in either liquefied or gaseous form. Some air separation plants build gas pipelines directly into large-volume contract customers.

Industrial gas production is electricity-intensive, accounting for 77% of the total industrial gas energy consumption. The process of separating air is primarily through a cryogenic technique. Air is compressed and cooled to temperatures below –200°F, then sent through turbines where additional energy is removed. After the air is liquefied, it is fractionated so that the component gases can be isolated. The gas products are then purified and reliquefied as required for shipment. The cryogenic technique is largely a mechanical vapor-compression process; consequently, motor drives account for about 87% of all electricity used by the industry.

21.8.2 Industrial Inorganic Chemicals

Industrial inorganic chemicals represents a diverse group of products, including pigments, acids, salts, activated carbons, phosphorus, and sulfur (chlorines, alkalis, and industrial gases are inorganic; they are treated as separate industry groupings). The inorganic chemical industry accounts for about 25% of all chemical products, supplying raw materials to the automotive, paper, packaging, pharmaceuticals, and paint industries.

21.8.3 Manufactured Fibers

The manufactured fiber industry converts polymers such as nylon, acrylonitrile, polyester, polyurethane, cellulose, and cellulose acetate into four primary types of fiber: staple fiber, filament, monofilament, and tow. These products are, in turn, used as raw materials by the textile, apparel, and cigarette industries. The textile industry is the largest industrial consumer of manufactured fibers. Machine-drive applications account for 54% of total electricity use. Energy accounts for between 10% and 18% of the total cost of goods.

In addition to being energy-intensive, the manufactured fiber industry requires highly reliable power. The nature of the manufacturing process makes any interruption in the product flow costly. Since fibers are continuously spun, any disturbance in machine speed or temperature can severely impact product quality. Unexpected downtime can create costly, time-consuming cleanups. Manufactured fiber facilities typically operate 24 h per day, 365 days per year to avoid the cost of restarting the processes. Consequently, power reliability and power quality are high priorities for the industry. Similarly, the industry uses a very conservative approach to investments in electrical equipment.

21.9 THE PULP-AND-PAPER INDUSTRY

21.9.1 Industry Organization

This industry is basically organized into (1) pulp mills and (2) paper and paperboard mills. The pulp-and-paper industry includes the manufacture of pulp from wood products, the production of paper and paperboard from wood pulp, and the conversion of bulk paper and board to finished products. It is sometimes considered part of the larger forest products industry, which includes the agricultural activities of tree farming, as well the conversion of the trees into lumber and wood products.

21.9.2 Pulp Mills

The pulping industry includes the manufacture of wood pulp, which is the input for the entire paper industry. Most of the pulp produced in the United States is for internal use in paper production, but 10% to 15% of the total pulp produced is sold on the open market. Types of pulp include bleached kraft, unbleached sulfite, and thermomechanical products. The "type" refers to the pulp manufacturing process. Pulp types are typically destined for a specific market, such as the use of bleached kraft pulp in higher-quality printing and writing papers. Chemical pulp accounts for approximately 90% of U.S. manufactured wood pulp and is based mostly on the kraft process. In this process, wood chips are mixed with a caustic sulfite cooking liquor at temperatures of 176°C. The wood is then separated into individual fibers with limited mechanical agitation. When the resulting pulp is washed, the waste liquid is processed in a recovery boiler. This boiler generates steam for both process use and power production, and the pulp reduction chemicals are also partially recovered for reuse. Mechanical pulping is very electrically intensive, and essentially applies heat and pressure to convert wood to pulp. Mechanical pulping processes use 1400 to 1700 kWh per ton of pulp produced.

21.9.3 Paper and Paperboard Mills

This industry includes the production of paper, paper products, and packaging materials. The paper industry companies are typically vertically integrated and incorporate most elements of forest products in their businesses. Environmental pressures have increased for the industry since the late 1980s, as in the treatment of the low levels of dioxins in paper-mill sludges. The "cluster rule," passed on April 15, 1998, regulates hazardous air pollutants and chlorinated organics (mostly phenolics). A focus of environmental concern is the continued use of bleaching technology that employs chlorine dioxide in the process. The chlorine dioxide is converted to sodium chlorate with the release of chlorine as a by-product. Mills are now introducing various control measures to deal with this issue.

Papermaking consists of stock preparation and formation of the final paper product in the paper machine. The primary step in stock preparation is refining the cellulose fibers in the pulp to facilitate good fiber-to-fiber bonding. The paper machine removes water from the pulp slurry, proceeding through processes of forming, pressing, and drying. To give an indication of the amount of water removal required in this process, approximately 100,000 gal of water is extracted for each ton of dry paper produced.

Stock preparation and refining requires 1 to 2 million Btu of thermal energy and 150 to 400 kWh of electricity per ton of paper processed. Paper refining requires 6 to 12 million Btu of thermal energy and 250 to 400 kWh of electricity per ton of paper processed. A newer technique of water removal—impulse drying—promises to improve overall efficiency with a lowering of thermal energy needs and an increase in electricity demand.

21.9.4 Power Distribution System

Mills typically use both 480- and 600-V systems for the lower voltage level, although 480 V is predominant. The mill may have its own substation, with distribution system higher voltages at 2400 to 15,000. Most large mills will have multiple 3-phase system voltages at levels including 480, 2400, 4160, or 13,800. Radial distributions are usually preferred because of their simplicity and lower cost.

21.10 DISTRIBUTED GENERATION

Distributed generation refers to small electric generating units located close to load centers. It encompasses onsite generation, self-generation, cogeneration, and any other small-scale power generation that is not considered central-station power. Distributed-generation projects are undertaken by end-use customers, electric utilities, gas utilities, power marketers, and other third parties.

The combination of electric utility deregulation, emerging generation technologies, and the growing worldwide demand for power has renewed interest in distributed generation. Electric utility industries are being deregulated in countries around the world. At the same time, new generation technologies such as fuel cells and microturbines are being commercialized, while conventional small-scale power generation technologies such as reciprocating engines and combustion turbines continue to make evolutionary improvements in performance. At the same time, worldwide demand for electric power has never been greater, and continues to grow, in both developed and developing countries. The combination of these three factors has led to a greatly increased interest in distributed generation.

21.10.1 Why Distributed Generation Is Used

Distributed generation is used for the following traditional applications:

- *Baseload power,* where the unit operates at a high-load factor throughout the year—usually the unit is operated as many hours as possible
- *Intermediate power,* where the unit operates during a specific period during the day, such as 9:00 A.M. to 6:00 P.M., during a specific seasonal or some other time period—usually 500 to 2000 h during the year
- *Peak shaving,* in which the unit operates only during the site's peak electricity demand or during peak electricity price periods—usually less than 500 h per year
- *Cogeneration* where the unit operates for a majority of the time during the year, and supplies both electricity and thermal energy to a site
- *Emergency/backup power,* where the unit is configured to supply power to the site in the event of a grid outage

In addition, a number of new niche applications for distributed generation have emerged, including

- *Premium power*—the unit supplies high power quality, free of interruptions, sags, surges, and spikes.
- *Green power*—the unit supplies power while generating low or zero emissions of pollution.
- *Power used as a hedge by power marketer*— in a deregulated electricity marketplace, strategically located small-scale power generation can be used as a source of power to fulfill contracts when the spot market price of electricity is high.
- *Remote power*—power is supplied where no distribution lines are available.

Various early adopters are currently using distributed generation for a number of reasons. Some use distributed generation purely on the basis of energy economics (distributed generation is cheaper for them than purchasing electricity from the grid). Others use small-scale onsite power generation because of the other benefits, such as increased reliability.

21.10.2 Distributed-Generation Technologies

A number of different technologies are available for distributed generation. Some of these are proven, mature technologies that have been used for decades. Others have been commercialized and have entered the marketplace only recently. Each is described below.

Reciprocating Engines. Otto and diesel cycle reciprocating engines were first developed in the late nineteenth century. They have gained widespread acceptance in almost every sector of the economy, and are used for applications varying from fractional-horsepower units used for small handheld tools to enormous 60-MW baseload electric power plants. Designs have been undergoing evolutionary changes and continue to improve in terms of efficiency, emissions, and other parameters.

For distributed-generation applications, there are two main types of reciprocating engines: Otto (spark ignition) and diesel (compression ignition). For most distributed-generation applications, these engines are of the four-stroke type, and use a piston that reciprocates in cylinders bore with the following cycles: (1) intake, (2) compression, (3) combustion, and (4) exhaust.

The piston starts at the top, the intake valve opens, and the piston moves down to let the engine take in a cylinder full of air and fuel during the intake stroke. Then the piston moves back up to compress this fuel/air mixture. Compression makes the explosion more powerful. When the piston reaches the top of its stroke, the spark plug emits a spark to ignite the fuel (in a diesel engine, the fuel is ignited by the compression alone). The fuel charge in the cylinder explodes, driving the piston down. Once the piston hits the bottom of its stroke the exhaust valve opens and the exhaust leaves the cylinder to go out the tailpipe. Now the engine is ready for the next cycle, so it takes in another charge of air and fuel.

Reciprocating engines are manufactured all over the world. They are currently being used in many distributed-generation applications, for primary power, intermediate power, peak shaving, and cogeneration. They are also used widely for emergency/backup power. They have a fairly low capital cost and reasonable electric efficiency, especially larger, low-speed units. For cogeneration applications, some low-quality steam can be produced, but most applications are for hot water.

Turbines. Combustion turbines are a widely used distributed-generation technology. They range in size from about 500 kW up to large utility-sized units of over 100 MW. Most of the smaller ones (less than 15 MW) are derived from aircraft engines; most of the larger ones are specifically designed for power generation. Combustion turbines are a proven technology that has been used for onsite power generation for many decades.

Combustion turbines consist of a compressor, a combustor, a turbine, and a generator. The compressors and turbines are typically multistage axial-flow designs, and somewhat resemble steam-turbine configurations.

Combustion turbines are used primarily for industrial and large commercial-sector cogeneration. They are also sometimes used for noncogeneration applications, and sometimes smaller units (usually less than 5 MW) are used for backup power. High-quality steam can be produced. Combustion turbines can be run on almost any gaseous or liquid fuel. Electric efficiency is usually poor (less than 35%, but overall efficiency in good cogeneration applications can be as high as 90%). At the time of writing, new combustion-turbine designs, utilizing technologies such as recuperators and ceramic components, with electric efficiencies of over 40%, are expected to enter the marketplace soon.

Microturbines. Microturbines are an emerging class of small-scale power generation technology in the 25- to 300-kW size range. The basic technology used in microturbines is derived from aircraft auxiliary power systems and diesel-engine turbochargers. A number of companies are currently field-testing demonstration units, and commercial deliveries have already begun. Microturbines consist of a compressor, a combustor, a turbine, and a generator. The compressors and turbines are typically radial-flow designs, and look much like automotive engine turbochargers. Most designs are single-shaft, and use a high-speed permanent-magnet generator that uses an inverter to produce ac power.

A number of different microturbine designs and configurations have been tested by manufacturers. The following are some fundamental design and configuration considerations:

Recuperation. Recuperators are air-to-air heat exchangers that use a microturbine's hot exhaust gases to heat the combustion inlet air after it has been compressed. Recuperators are key to the electrical efficiency of microturbines.

Single-shaft with high-speed permanent-magnet generator and inverter versus split-shaft with reduction gearbox and 60-Hz induction generator. Two different ways of producing ac power with microturbines are currently being evaluated in the demonstration units. The first and more common uses a single shaft with a high-speed permanent-magnet generator spinning at the same speed as the turbine. This generator produces very high-frequency ac power that must be converted to 60 Hz using an inverter. The second design uses a two-shaft configuration with a reduction gearbox and a 2-pole, 3600-r/min induction generator that directly produces 60 Hz of power.

Air-bearing versus oil-lubricated. Microturbines are high-speed (≥40,000 r/min) rotating equipment that require high-reliability bearing systems. Two different configurations are currently being used. The first uses air bearings with a compliant foil system that requires no oil lubricant. The second uses a pressurized lube-oil system with a pump, similar to that used by an automobile engine.

Fuel-cell hybrids. Power generation systems utilizing fuel cells combined with microturbines are also being developed by several manufacturers. These systems typically run the hot gas produced by fuel cells through a microturbine to generate additional electricity. Hybrid system are predicted to have exceptionally high electric efficiencies (≥60%).

Gasifiers. Gasifiers produce gaseous fuel from solids, such as coal and biomass. Small-scale gasifiers for use with microturbines are still in the development stage. Gasifiers could help microturbines gain wider acceptance, especially in international markets.

With recuperation, microturbine efficiency is projected in the 26% to 30% (LHV) range. The overall efficiency is projected at 65%. Overall efficiencies may reach 85%, similar to larger turbines. Microturbines can run on natural gas, diesel, gasoline, kerosene, naphtha, methanol, ethanol, alcohol, propane, JP-8 Flair Gas, and other gases with heating values over 500 Btu/lb.

Fuel Cells. Fuel cells are an emerging class of small-scale power generation technology in the ≥25- to 1000-kW size range. The first fuel cells was developed in 1839 by Sir William Grove, although they were not used as practical generators of electricity until the 1960s, when the U.S. space program chose fuel cells for the *Gemini* and *Apollo* spacecraft.

There are a number of fuel-cell types and configurations, but they all use the same basic principle. A fuel cell consists of two electrodes sandwiched around an electrolyte. Oxygen passes over one electrode and hydrogen over the other, generating electricity, water, and heat. Hydrogen fuel is fed into the anode of the fuel cell. Oxygen (or air) enters the fuel cell through the cathode. Encouraged by a catalyst, the hydrogen atom splits into a proton and an electron, which take different paths to the cathode. The proton passes through the electrolyte. The electrons create a separate current that can be utilized before they return to the cathode to be reunited with the hydrogen and oxygen in a molecule of water. A fuel-cell system usually includes a "fuel reformer" that can utilize the hydrogen from any hydrocarbon fuel—from natural gas to methanol, and even gasoline.

Other Technologies. Other technologies that fall under the distributed-generation umbrella include renewables such as photovoltaic (PV) and wind power. PV technologies convert sunlight directly to electric power. Current PV technologies have very high capital costs, but they continue to improve. Wind-power systems use turbines turned by wind to provide torque to a shaft that turns a generator. These technologies also continue to improve, and now can provide electricity at competitive rates, although they obviously operate only under the proper atmospheric conditions.

21.11 ELECTRIC MELTING

21.11.1 Process Overview

Electric Melting Applications. Electricity has been used extensively for industrial process melting applications since the early 1960s, including iron, steel, nonferrous metals, and glass. The most important applications of electric melting are

1. Melting and refining of steel
2. Melting, holding, and casting of most metals, including iron, steel, and nonferrous metals
3. Melting, refining, forming, and annealing of glass

Glass Melting. Electricity has been used in glass production to melt, refine, form, and anneal glass to the desired properties for use in containers, flat glass, pressed and blown glass, and glass fibers. Applications for electrically melting glass were limited until the mid-1950s, and since then its use in either all-electric furnaces or electrically boosted melting has grown substantially.

Technology Applications. The following sections discuss four technologies used for electric melting applications.

Metal melting includes, in addition to the change of state, the further heating of the metal to a specified temperature, known as the "pouring temperature," if the metal is to be poured into a mold or the "working temperature" if the metal is to be used for coating, as in galvanizing, or as a liquid heating bath.

Table 21-1 refers to melting service. An alloy belongs in a group of this table to which its major component belongs. The melting-temperature range of an alloy must be obtained from the constitution diagram of a system to which the alloy belongs.

TABLE 21-1 Melting Points of Metals

Metal	Melting point °C	Melting point °F
Group 1		
Tin	232	450
Bismuth	271	520
Cadmium	321	610
Lead	327	621
Zinc	420	788
Antimony	630	1166
Magnesium	651	1204
Aluminum	659	1218
Group 2		
Silver	961	1761
Copper	1083	1981
Nickel	1452	2646
Cobalt	1480	2696
Iron	1530	2786

The pouring temperatures of nonferrous metals and alloys range from 100 to 200°C (180 to 360°F) above the melting point. They vary for each metal or alloy with the type of load and the size and type of casting according to the purpose for which the casting is to be used.

Volatilization occurs. In melting metals there is a certain loss by volatilization. This loss of metal is usually negligible except for charges which contain a high percentage of zinc. Another consideration is the poisonous nature of zinc fumes.

Stirring an alloy while it is in the molten state is often necessary to prevent segregation. Stirring is also an aid in bringing about equalization of temperature and is a safeguard against overheating in the surface of the molten mass.

Automatic stirring can be obtained with suitable induction coils or by the use of an induction furnace.

21.11.2 Melting Pots

The melting of the soft metals, or group 1 (see Table 21-1), and their alloys is within the range of the 80 Ni-20 Cr alloy-resistor heating equipment. Containers for these materials are designated as melting pots, solder pots, lead pots, galvanizing kettles, tanks, etc.

Melting pots are made of cast iron or steel and of various nonferrous alloys. The selection of the material for the pot depends on the temperature of the molten metal and the possible reactions of that metal with the material of the pot.

Open-top pots are used in which the molten metal is removed by dipping or pumping and into which metal is dipped for coating. A greater depth than is necessary for the service is desirable as an aid to temperature recovery when cold metal is added.

Closed-top pots are advantageous in reducing the heat loss from the molten-metal surface and in reducing oxidation. Also, space can be provided for a protective atmosphere. Discharge of the molten metal by gas pressure, for example, steam, is incorporated in some designs.

The heat insulation of a metal melting pot should be sufficient to limit the outside surface temperature to a safe value. Otherwise, insulation is a matter of economy. As a rule, a layer of refractory material is unnecessary for temperatures below 600°C (1112°F). A refractory layer is useful as an aid in maintaining an even metal temperature.

Heat loss occurs. The rates of heat loss can be reduced by a layer of insulating material in granular or gaseous form, for example, charcoal or diatomite on the surface of the metal or sulfur dioxide. This layer also reduces oxidation of the metal. When using gaseous layers, adequate ventilation must be provided.

A *cast-in unit* is a self-contained heating unit embedded by casting in a mass of gray iron. This unit is designed for metal-melting services up to 510°C (950°F). The unit is applicable to the melting of all the soft metals with the exception of aluminum and zinc, exceptions because of their alloying action with iron. The rating of the unit should be on the open-air basis.

External heating units supplied to a melting pot should be rated on the basis of 20 to 20 W/in^2 of the side surface of the container. This design of melting pot is suitable for molten-metal temperatures up to 898°C (1650°F).

The temperature-regulating equipment of melting pots is the same as used for resistor ovens and furnaces. As a rule, small melting pots—say, 5 kW and below—have only manual control by a two- or three-point switch.

The melting rate of a melting pot is measured by the time required for the charge of molten metal to regain the pouring or working temperature after the addition of a quantity of cold metal. For example, if 100 lb of cold metal is added and the metal in the pot regains its temperature within 10 min, the melting rate for that metal is $100 \times 6 = 600$ lb/h.

The kilowatts rating of a melting pot is based on the required rate of melting a given metal or alloy. For melting quantities of metal, for example, for castings, the kilowatts rating should be no less than the rate of heat input to the metal plus the rate of heat loss. For melting for coating work, for example, galvanizing, the kilowatts capacity needed is the sum of the capacities required for melting, for heating the base material, and for the rate of heat loss. As a rule, some additional capacity is installed to accelerate heating up and to prevent too large a drop in temperature when cold metal is added. The operating efficiency of the melting range is not affected by its kilowatts rating.

Metals of group 2 and their alloys require either the arc furnace or the induction furnace for melting. Brass and steel are the major alloys of this group.

21.11.3 Arc Furnaces

The two types of arc furnace in common use are (1) the 3-phase furnace and (2) the single-phase furnace. The general field of the 3-phase furnace is the production of alloy steels; that of the single-phase furnace, the production of nonferrous alloys. Both types of furnace can be used for the manufacture of high-quality gray-iron castings. However, the use of large-size induction coreless and core-type furnaces may be preferable.

Three-Phase Arc Furnace. Standard sizes of arc furnaces range from 250 to 80,000 kVA; loading range, 500 lb to 250 tons. Sizes 1000 to 5000 kVA predominate.

The chamber is a steel bowl with a refractory lining. The hearth is a shallow bowl formed in the bottom lining. The roof is a removable dome-shape refractory structure carried on a steel roof ring. The roof has three round ports in equilateral triangular arrangement through which vertical carbon or graphite electrodes travel. Each electrode is carried on a winch-and-rope system, motor-driven.

Refractories. The chemical nature of the slag, acid or basic, determines the required chemical nature of the lining of the hearth and sidewall of the chamber up to a few inches above the top surface line of the slag, that is, an acid refractory (silica) for acid slags, and a basic refractory (magnesia) for basic slags. The roof is usually made of silica brick. Silica has a tendency to spoil during heating and cooling, and furnaces which are in intermittent use often have roofs made of fire-clay brick.

Temperature. The operating temperature of the chamber is limited by the softening point of the refractory, particularly that of the roof, where there is a concentration of heat. A refractory material can be operated with the temperature of its inner face close to its softening point provided that the outer face is exposed to the open air, thus permitting a flow of heat through the refractory body. A temperature gradient is thus established in the refractory so that if its thickness is correctly related to its thermal conductivity the mean temperature of the refractory body will not be high enough to impair its strength materially.

The temperature of molten steels is around 1600°C (2912°F). The melting point of silica is 1713°C (3115°F), but the softening point of a silica refractory is somewhat lower because of impurities in the refractory body. Hence, in a steel melting furnace the temperature of the inner face of the refractory lining is too high to permit the use of heat insulation. Even a thick coat of dust on the roof of the melting furnace is undesirable.

The designation for a 3-phase arc furnace may be given in terms of the holding capacity, the shell diameter, the pouring capacity, the melting rate, or a combination of these. A given diameter of shell can be attached to a range of ratings by varying the thickness of the refractory linings. Sizes are given in Table 21-2.

Charges. The 3-phase arc furnace is primarily a scrap-metal-conversion unit. The two types of furnace with respect to the method of charging are (1) the door-charge type and (2) the top-charge

TABLE 21-2 Representative Sizes of 3-Phase Arc Furnaces in General Use

Shell diameter	Normal charge, lb	Transformer rating, kVA	Single slag heats, lb/h
4 ft	800–1,000	250–350	500
4 ft 6 in	1,200–1,500	350–500	900
5 ft	1,500–2,000	500–750	1,300
6 ft	3,000–4,000	750–1000	2,000
7 ft	5,000–6,000	1000–1500	3,000
8 ft	7,000–9,000	1500–2000	4,500
9 ft	10,000–12,000	2000–3000	6,000
10 ft	16,000–20,000	2500–3000	10,000

Note: 1 ft = 0.3048 m; 1 lb = 0.4536 kg.

TABLE 21-3 Approximate Current-Carrying Capacities of Graphite Electrodes for Arc Furnaces

Nominal diameter, in	Amperes	Nominal diameter, in	Amperes
2	600–1000	9	6,400–10,800
2½	800–1500	10	7,800–12,500
3	1200–2100	12	11,300–17,000
4	1800–3000	14	15,400–21,500
5½	2300–4100	16	20,100–26,100
6	3100–5400	17	22,700–27,400
7	4200–6900	18	25,500–30,500
8	5500–9000	20	28,300–34,600

Note: 1 in = 25.4 mm.

type. Depending on the character of the scrap, hand charging and chute charging are the usual methods for small furnaces. Large furnace installations are often equipped with side-door-charging machines. Top charging is favored for medium-size furnaces. In this method the roof of the furnace is removed, and a complete charge is placed in the chamber by a drop bucket handled by an overhead crane. This is both a time-saving and a labor-saving method. The charging time is only a few minutes, for example, a reduction from 30 to 5 min. Top charging has the other advantage of a full chamber and a lower heat loss during the charging.

Some 3-phase arc furnaces are used for refining service only. Molten metal from an open-hearth furnace, Bessemer converter, or cupola is the charge.

The weight of scrap metal varies with the degree of its subdivision. The weight of charge that can be placed in a given furnace depends on the kind of scrap. If sufficient scrap metal cannot be placed in the furnace initially to form the weight of molten metal desired, additional quantities can be added later in the heat cycle. This practice affects adversely to some extent both the operating efficiency and the consumption of electrodes.

Electrodes. The arc in each phase is maintained between the tip of the electrode of that phase and the charge (bath after the molten state is reached). The charge thus serves as a common electrode for the three arcs and makes a connection of the 3-phase circuit at that point. The designation "direct-arc furnace" refers to this arrangement.

The trend is toward the general use of graphite electrodes. Carbon electrodes are preferred in some cases. Standard sizes in the corresponding current ratings are given in Tables 21-3 and 21-4.

The consumption of electrodes is caused largely by volatilization and burning. There is some breakage. Graphite begins to oxidize at about 600°C; carbon, at 400°C. Under average conditions the consumption of graphite electrodes is about one-half that of the carbon electrodes. Average values for melting service, pounds of electrodes per ton of metal melted, are graphite 4 to 10; carbon 8 to 15. The corresponding consumption in melting refining service is about 10 lb for graphite and 18 lb for carbon.

Voltamperage. The voltampere characteristic of the arc is negative and a stabilizing element is necessary for circuit stability. Reactance also serves to limit the current in the circuit when an

TABLE 21-4 Approximate Current-Carrying Capacities of Carbon Electrodes for Arc Furnaces

Nominal diameter, in	Amperes	Nominal diameter, in	Amperes
8	2,000–3,000	20	11,000–17,300
10	3,000–4,800	24	15,800–24,800
12	4,500–6,800	30	24,700–35,300
14	5,400–8,500	35	28,800–38,400
17	7,900–12,500	40	37,700–50,200

Note: 1 in = 25.4 mm.

electrode touches the charge. This reactance is a total reactance of the circuit from the furnace terminals to the point in the power system where the voltage is held constant. Thus a furnace at the end of a long feeder is a different problem from a furnace installed adjacent to a large substation.

The operation of an arc furnace is dependent on the stabilizing element of the circuit only to the extent of ensuring continuity of operation. The limitation of current fluctuations is a problem of power service and is individual for each location. The resistance of the circuit is also a factor, and the actual value of the short-circuit current will be less than that indicated.

Circuit Characteristics. The arc-furnace circuit (containing resistance and reactance) is operated at constant voltage and supplies the unity power-factor load, the arc or arcs. The maximum power of the circuit occurs at 0.707 power factor. The maximum power in the arc occurs at a higher power factor of the circuit, a value dependent on the constant of the circuit.

Electrical Apparatus. The rating of the electrical equipment of a 3-phase arc-furnace installation varies for a given size furnace for the class of service and in some cases according to the power-service conditions. The electrical equipment includes

1. A variable-ratio power transformer
2. Reactors if required
3. An automatic current regulator
4. A control panel for the operator
5. Electrode motors and tilting motors
6. A main-line circuit breaker and disconnecting switches

Transformers. The features which distinguish the arc-furnace transformer from the conventional power transformer are (1) individual service, (2) no requirement of regulation, and (3) a wide range of comparatively low secondary voltages and correspondingly high secondary currents.

Reactance. There are no criteria for stability in current limitations in arc-furnace circuits and hence no standard values of reactance in these circuits. As a rule, 40% to 60% reactance is satisfactory.

The inherent reactance in the circuit of a large furnace—500 kVA and larger—may be, and usually is, sufficient for the need. As the secondary voltage is fixed by conditions other than the kVA rating of the circuit, the smaller installations require more or less supplemental reactance.

The normal reactance of 60-Hz furnace transformers ranges from 5% to 7%. Reactance values higher or lower in each case, within certain limits, than the range noted can be obtained by design but may entail a sacrifice one way or another in the design of the transformer. Hence it is considered better practice to use a normal design of transformer and to add supplemental reactance, if needed, by reactors.

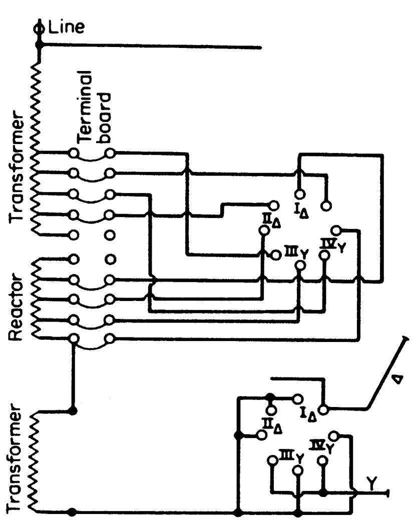


FIGURE 21-6 Diagram for four operating voltages, one phase only of a 3-phase transformer.

Reactors should have a number of taps for adjustment after installation. The transformer taps and reactor taps are connected into a common terminal board arranged so that any combination of transformer taps and reactor taps can be made for each of the selected operating voltages.

The diagram in Fig. 21-6 illustrates a tap arrangement and switching arrangement for operating voltages and the use of reactor windings for both the delta and the Y connection.

An example is the connection of one installation given in the table.

Automatic Current Regulator. A change in current causes a change in the power of an arc-furnace circuit. Within the limits of circuit stability the current with a given applied voltage can be changed by changing the length of the arc. This is the principle of the power regulation of arc furnaces.

The intermittent-type current regulator, which operates within preset limits of current variation, has been superseded almost entirely by the continuous-type current regulator.

A simplified diagram, one phase only of the continuous-type regulator, is shown in Fig. 21-7. The principle of operation is the opposition of a voltage B derived from the circuit of the arc by a reference voltage A. The resultant value of these two voltages determines the polarity of the generator that drives the electrode motor. Thus the length of the arc—and correspondingly the current in the arc—is maintained at a predetermined value.

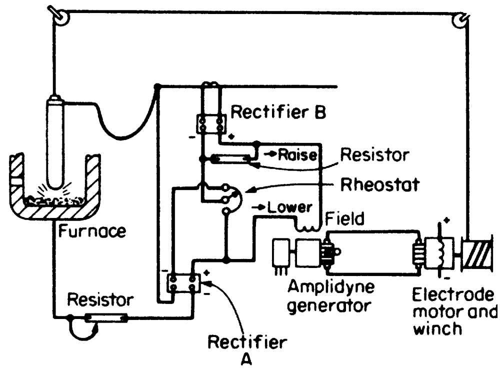


FIGURE 21-7 Simplified circuit diagram of continuous-type automatic current regulator for a 3-phase arc furnace (single phase only).

Each phase in the 3-phase circuit is regulated independently. However, because of the common electrode, the charge, or both, the three elements of a 3-phase regulator work together for the maintenance of equal currents in the three circuits of the power system.

The automatic regulator performs three other functions:

1. The feeding of the electrodes at the rate of consumption
2. The removal of partial short circuits caused by the electrodes coming into contact with the charge
3. The protection of the equipment in case of a failure of the power supply

Operator's Panel. The standard equipment consists of three ammeters, a polyphase wattmeter, a voltmeter, and the necessary rheostats and switches for the operation of the furnace.

The Electrode Drive. This is a reversing service—rapid at certain times—and a motor with a low WR^2 effect is desirable.

The Main-Line Circuit Breaker. This serves both as a protective device and as a switch. The switching service rate is many times per day. In normal operation the arc circuit is opened by raising the electrodes so that the circuit breaker opens the magnetizing part of the power transformer.

All changing of taps is done with a main-line circuit breaker open (no-load tap changing). The tap-changing switch is interlocked with the circuit breaker to prevent incorrect operation. As a rule the tap-changing switch is mounted inside the tank in 3-phase transformers and outside when three single-phase units are used. The tap-changing switch can be motor-operated or hand-operated; the former is the more general practice.

Single-Phase Arc Furnace. The most common single-phase arc furnace is the automatic rocking furnace. This furnace is used extensively for melting both ferrous and nonferrous metals and alloys. Standard sizes extend up to and include 600-kW ratings for melting 4000 lb of cold steel scrap in 90 min.

The load characteristics of a single-phase arc furnace are similar to those of 3-phase arc furnaces. However, as there is no arc between an electrode and the charge, the initial performance of the single-phase furnace is somewhat better than that of the 3-phase arc furnace. The average power factor of the single-phase furnace is 70% to 80%.

Electrical equipment for single-phase arc furnaces is similar to that for 3-phase arc furnaces. Usually only one operating voltage is used.

Gray iron with uniform structure and high engineering properties, namely, tensile, bending, shearing, and impact strength, is produced in arc furnaces. The method of production involves superheating the iron after melting to about 1600°C (2912°F) and holding it at the elevated temperature for a brief period of time. Gray irons with the tensile strength of 40,000 lb/in^2 are produced regularly by this method.

21.11.4 Induction Furnaces

Two types of metal-melting furnace that embody the induction principle are the coreless furnace, which can be used in operation from frequencies of main-line 60 Hz up to many hundreds of kilocycles (the frequency generally depends on the size of the furnace) and the channel furnace, most generally operated from main-line frequencies, principally 60 Hz.

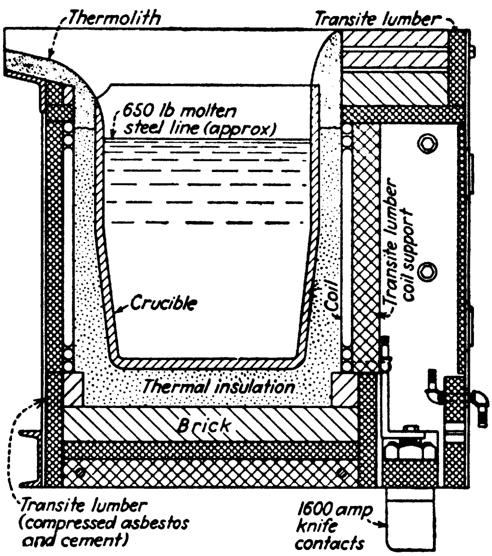


FIGURE 21-8 Coreless induction furnace.

The Coreless Induction Furnace. The general design of this furnace is shown in Fig. 21-8. The assembly consists of three main parts: (1) the primary coil, (2) the refractory container, and (3) the frame, which includes supports and tilting mechanism.

The distinctive feature of this furnace in common with other assemblies for induction heating is the absence of a continuous iron path for the magnetic flux. However, iron laminations are frequently used on the larger sizes, and particularly for line-frequency furnaces, to reduce the reactance for the flux on the back or outside of the furnace. Another feature for comparison with other types of melting furnace is the small quantity of refractory material in the construction.

Standard preformed crucibles are used for the smaller furnaces, up to about 500-lb holding capacity. Over 500-lb capacity rammed linings are employed rather than preformed crucibles. The preformed crucibles are generally used for nonferrous melting.

Standard sizes for steel melting furnaces are 50 up to 60,000 lb. These are often listed in the ratings of 100-lb capacity, 200, 300, 600, 1000, 1500, 2000, 4000, 8000, 10,000, 15,000, 20,000, and 40,000-lb capacity, plus the various intermediate sizes.

Features of operation peculiar to this type of melting furnace are that

1. The refractory container makes necessary a large air gap (loose coupling), with consequent low power factor, 20% to 30%.
2. The charge is cold scrap metal. Thus, initially, the secondary circuit is a current path through a variety of shapes and dimensions of pieces, and contact resistance is a large part of the total resistance. The charge becomes homogeneous as the metal melts, shrinks in height, and thus decreases the coupling of the circuit. At this time, cold metal may or may not be added to the charge. It is a common practice after the first melt to maintain a molten heel in the furnace to expedite the melting of the second load.

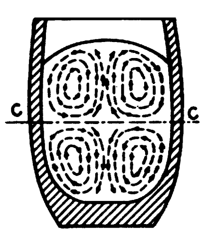


FIGURE 21-9 The stirring effect in a mass of molten metal being heated by induced currents.

3. With charges of magnetic material the effect of the magnetic property is pronounced at the start of the heat cycle.
4. Stirring of the molten metal, as indicated by Fig. 21-9, is characteristic of this furnace. This movement of the bath has much metallurgical significance in the production of homogeneous alloys. The movement is often desirable and can be regulated to a certain extent by the frequency selected for the size furnace involved.

The characteristics of charges vary widely, but the general practice is to vary the capacitance on the circuit during the heat cycle to maintain approximately unity power factor.

Frequency. The primary technical factor in the selection of frequency for a metal-melting furnace is the desired degree of stirring of the molten metal for the size of furnace required. This stirring effect is proportional to the square of the ampere-turns and inversely proportional to the frequency. For a given power value, the current decreases and the voltage increases with increase of frequency.

A second consideration is the fineness of the scrap metal of the charge. If the pieces are very small, there may be difficulty in starting the melting of a cold charge. The frequency must be high enough in a given case to give an electrical efficiency high enough for starting. These conditions and the economics of the service early led to the adoption in this country of 960 Hz for steel-melting furnaces, 100 kW and above, and 3000 Hz for smaller furnaces. As a rule these frequencies are also suitable for melting nonferrous charges. Various frequencies are used for laboratory furnaces. In general industrial practice the trend is to use 60 Hz where large-tonnage furnaces are being used.

Service. The coreless induction furnace is primarily a metal-melting unit. An important use of this furnace is the production of carbon ferrous alloys. The refining of steel, that is, the removal of phosphorus and sulfur, in this furnace has not been developed to a fine art. The deoxidation of the melt, for example, by the addition of aluminum just before pouring, is not here classed as refining. Various special services are duplexing steel, vacuum melting, heating of charges of nonconducting material (with or without melting) by the use of conducting crucibles, etc.

Performance. The melting rate of a furnace, and consequently the power input, is determined in each case by the rate at which the molten metal can be used. Heat cycles of 1/2 h for small furnaces, 1 h for medium-sized furnaces, and 1 1/2 to 2 h or more for large furnaces represent typical practice.

Energy Consumption. This varies for a given material with the size of the furnace, the melting rate, and the idle time between heats. Representative values are 330 kWh/ton (2000 lb) for copper 500 kWh/ton for gray iron, and 600 kWh/ton (2000 lb) for steel. These figures are lower with larger furnaces in continuous operation and somewhat higher with small units and infrequent service.

The Channel (Core-Type) Furnace. The basic feature of this furnace is a single-turn loop of molten metal below and connected to the bath serving as a secondary circuit. A conventional laminated steel core and primary winding complete the transformer feature.

The heat is developed in the loop of molten metal below the body of the charge. Electrodynamic action causes motion of the metal in the loop as indicated. This movement serves to transfer the heat and so stir the molten metal in the chamber above the loop.

Starting the furnace requires sufficient molten metal to close the secondary circuit. In changing from the metal or alloy to another, the furnace must be emptied. For day-to-day melting of the same metal or alloy—the usual practice—the charge, or a portion of the charge, is held molten overnight by using a below-normal voltage.

The V-shaped loop furnace is designed for melting the heavy nonferrous metals and alloys. It is used widely for melting brass. With these metals the slags, or nonmetallic particles, in the melt are much lighter than the metal, so that they tend to float on the surface of the bath and do not interfere with the circulation of the metal in the loop of the secondary circuit.

Standard sizes on these furnaces range from 60 to 1000 kW on single-phase, 60 Hz in the United States, and for standard voltages up to 600 V inclusive.

A design of the channel furnace for melting aluminum and its alloys embodies two or more vertical channels below the bath and connected to the bottom by horizontal channels of larger cross section. The slag particles of the lighter metals have about the same specific gravity or are heavier than the molten metal. Such particles tend to accumulate in the channel of the secondary circuit. The design of the channel noted above provides for its easy cleaning from time to time while the furnace is in operation.

Power and power factor of the channel furnace with constant applied voltage vary with the resistivity of the metal in the loop that forms the secondary circuit.

21.11.5 Resistance Furnaces

Resistance furnaces have a temperature range of 1500°C (2732°F) and above. This type of furnace is generally an open-top heating chamber with electrodes—movable or fixed—buried in the charge. Variations are furnaces with closed tops, furnaces with a resistor buried in the charge, etc.

The general service is heating charges of a refractory nature to bring about chemical reactions or changes in the physical structure of a material. Where the product is obtained by a chemical reaction, for example, the reduction of a metallic oxide, the term *smelting furnace* applies.

The length and cross-sectional area of the path of the current through the charge are proportioned to suit the power characteristic of the material to be heated. Where the buried resistor is used, its design is based on similar considerations. In all cases the load is resistance, that is, the I^2R effect. In some cases where the charge forms a resistor, there may be arcing along the path of the current.

Glass applications require conventional fuel firing initially to melt the glass. Once molten, glass conducts electrically and can be processed using bottom or side electrodes.

There are no standard designs of resistance furnaces. Each application is an individual undertaking, and the furnace is assembled in place. The simplicity of the construction permits a wide range of designs and much latitude in dimensions.

Resistance furnaces are unity-power-factor loads, and the circuit is stable; that is, reactance is not needed as with other furnaces. However, there are reactances and resistances in the furnace conductors, and the circuit characteristics are the same as for arc-furnace circuits. The voltage is low, generally less than 250 V. The corresponding high-current values introduce problems of inductance in the leads to the furnace, particularly with 60-Hz power supply.

Furnaces with Movable Electrodes. There are stationary, vertical, cylindrical, or rectangular structures, single- or 3-phase (see Fig. 21-10). The latter is more common.

The material of the charge, granular in form, is fed, more or less continuously, into open-top furnaces or through chutes into closed furnaces. The product, if molten, is discharged by tapping through a side, usually near the bottom. Vapor products pass out at the top or through the side openings in the chamber. If the product is solid, for example, fused magnesia, the batch method is used and the furnace is disassembled for removal of the product.

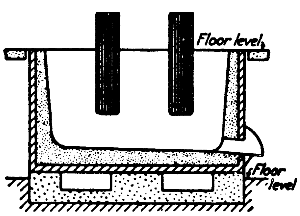


FIGURE 21-10 Open-type resistance furnace with movable electrodes (three electrodes for a 3-phase furnace).

The regulation of movable-electrode furnaces is accomplished by electrode movements to hold constant current, as with arc furnaces. The load is fairly uniform, and the regulating duty is light. Either dc motors or ac motors can be used to drive the electrodes. Direct-current motors are more often employed because of the ease of speed adjustment and the superiority of dynamic braking.

Continuous operation at one voltage is general practice. Transformers for these furnaces have a number of taps in the primary windings for the selection of the secondary voltage to suit operating conditions, which may change from time to time.

Batch Furnaces with Fixed Electrodes. A horizontal rectangular chamber with an electrode at each end is typical construction where the current is passed through the charge.

Refractory materials have negative resistivity versus temperature coefficients. Hence, to maintain constant power in a batch furnace as the temperature of the charge rises, the applied voltage must be varied as expressed by Eq. (21-12) or (21-13). Accordingly the regulation of furnaces of this type is by voltage adjustments, either load-ratio control or no-load tap changing, the latter with or without an induction regulator for intermediate voltages. Generally manual operation is sufficient. Constant power is not the rule. In many cases the power input is varied to correspond to a prescribed heat cycle.

Some of the more common uses of resistance furnaces are for calcium carbide, ferroalloy, ferrosilicon, ferromanganese, ferrochromium silicomanganese, ferrotungsten, ferromolybdenum, silicone oxycarbide, graphite, fused alumina, fused magnesia, fused silica, zinc oxide, and pig-iron applications.

21.12 ELECTRIC HEATING

21.12.1 Principles of Heating

Material is heated for processing purposes either (1) in the open air or (2) protected from the air in a vacuum or a special gas chamber. In comparative terms, in open air the rate of heating is fast and the time short. With enclosures the rate of heating depends on the source. It may be fast; generally it is slow, however, and the time is long. The basic consideration in each case is the correlation of temperature and the time, dictated by the material to be heated and the results desired.

Electric heating has certain unique characteristics:

1. The precision of electric control is extended to the transfer of heat. Uniformity of temperature within a relatively narrow limit is readily attained.
2. Its development does not involve combustion.
3. There is no upper limit to the temperature obtainable except the ability of the materials to withstand heat.

The collateral merits of electric heat vary in value with the conditions in each case. The most important of these are

1. Application at the precise point needed
2. Flexibility, includes easy subdivision, freedom of location, and general adaptability
3. Good working conditions, cleanliness, quietness, ambient temperature minimally affected, etc.
4. Fast response
5. Safety

Heat-Process Engineering. The heating method for a particular process may be selected after consideration of two basic principles of industrial heating: (1) temperature and (2) mode of heat transfer. This narrows considerably the field from which a choice of the heating method should be made. Generally there is more than one possible method. Economic factors must then be taken into account. This, in addition to the relative capital, maintenance, and energy cost, includes a study of the collateral merits of methods that always have a bearing on the overall cost of production.

A convenient temperature classification is

Low temperature—up to 400°C (752°F)

Medium temperatures—400 to 1150°C (752 to 2102°F)

High temperatures—beyond 1150°C (2102°F) (see section on electric melting)

Heat energy is transferred by conduction, radiation, convection, high-density electromagnetic force, and concentrated electrostatic fields. The transfer can be obtained with a combination of these methods.

Heat conduction can take place in all three states of matter, that is, in solids, in liquids, and in gases. Thermal conductivity is defined as the quantity of heat which flows in unit time through unit area of a plate of unit thickness having unit difference of temperature between its faces. There is no generally accepted combination of units for expressing thermal conductivity. Some of the more common combinations of units for this value are:

$$k = \text{Btu/ft}^2 \cdot \text{h for 1-in length of path/°F (British units)}$$

$$= \text{cal/cm}^2 \cdot \text{s for 1-cm length of path/°C (cgs units)}$$

$$= \text{W/(in}^2) \cdot \text{(1-in length of path)} \cdot \text{(°C)}$$

TABLE 21-5 Thermal Conductivities k^*

Solids at or near room temperature	
Silver	2824
Copper	2661
Aluminum	1463
Zinc	769
Iron	467
Nickel	412
Marble	21
Ice	15
Porcelain	7
Chalk	7
Glass	5
Liquids at temperatures noted	
Water, 20°C	4.15
Glycerin, 9–15°C	1.85
Ethyl alcohol, 20°C	1.23
Petroleum, 13°C	1.03
Gases at temperatures noted	
Hydrogen, 100°C	1.00
Methane, 7–8°C	0.188
Air, 0°C	0.165
Ammonia gas, 0°C	0.130
Carbon dioxide, 0°C	0.090

*British units; divide by 2903 to obtain corresponding cgs values of *k*.

Values of *k* for various materials as published are not based on a common method of determination. Hence, agreement in thermal-conductivity data for a given material cannot be expected.

The thermal conductivity of a material is affected by temperature, in some cases increasing, in other cases decreasing with rising temperature. Values representative of the general order of the thermal conductivities of ordinary temperatures of the three states of matter are given in Table 21-5.

An arbitrary standard by which to judge the value of a material as a nonconductor of heat (heat insulator) is $k = 1$ (British unit).

The law of thermal conduction for a constant-temperature gradient is expressed thus:

$$h_c = \frac{Ak_m t'}{l} \tag{21-1}$$

where h_c = rate of heat transfer by conduction, A = area of cross section of path of heat flow, t' = temperature gradient, l = length of path, and k_m = mean value of thermal conductivity of material in which the temperature gradient is established for a given range of temperature.

Shape Factor. Equation (21-1) as written is applicable only to paths of uniform cross section. If the area of the cross section of the path of heat flow varies along its length, the equation is applicable if the logarithmic mean area of the path is used.

For boundary surfaces that are not smooth curves, a shape factor *S* must be substituted for the term *A/l* in Eq. (21-1); thus

$$h_c = Sk_m t' \tag{21-2}$$

Two of the equations for the shape factor *S* are

1. Rectangular enclosures (Fig. 21-20):

$$S = \frac{A}{t} + 0.54\Sigma l + 1.20t \tag{21-3}$$

where A = inside surface area, t = thickness of wall, and Σl = sum of length of all inside edges.

2. Cylindrical enclosures (Fig. 21-21):

$$S = \frac{A}{t} + 0.54\Sigma l + \frac{2\pi h}{2.30 \log (b/a)} \qquad (21\text{-}4)$$

where A = inside area of top and bottom surfaces, t = thickness of wall, Σl = sum of lengths of inside top and bottom edges, h = inside height, b = outside diameter, and a = inside diameter.

Radiation. Radiation is a surface phenomenon. Surfaces emit and absorb radiant energy in various degrees, depending on the nature of the surface.

Average values of radiation and absorption coefficients are given in Table 21-6.

Solids for the most part are opaque to radiation; a notable exception is quartz. Liquids transmit radiation to some extent. Only two gases, water vapor and carbon dioxide, absorb radiation to any marked degree.

The rate of heat transfer by radiation under the conditions usual in heating practice is

$$h_r = \epsilon \times 3.68(T_0^4 - T_a^4) \times 10^{-11} \qquad \text{(Celsius scale)} \qquad (21\text{-}5)$$

$$h_r = \epsilon \times 0.35(T_0^4 - T_a^4) \times 10^{-11} \qquad \text{(Fahrenheit scale)} \qquad (21\text{-}6)$$

where h_r = watts per square inch, ϵ = absorption coefficient of receiving surface, T_0 = absolute temperature of emitting surface, and T_a = absolute temperature of absorbing surface.

Natural Convection. Natural convection is a transfer of heat to or from a surface by the movement of a fluid or gas when this movement is caused solely by a difference in fluid or gas density.

For natural convection in air

$$h_b = 11.70 \times t \times 1.25 \times 10^{-4} \qquad \text{(Celsius scale)} \qquad (21\text{-}7)$$

$$h_b = 6.50 \times t \times 1.25 \times 10^{-4} \qquad \text{(Fahrenheit scale)} \qquad (21\text{-}8)$$

TABLE 21-6 Average Values of Radiation (and Absorption) Coefficients

Material	ϵ radiant-energy coefficient
Blackbody	1.0
Lampblack	0.95
Asbestos board	0.93
Steel, oxidized	0.79
Copper, oxidized	0.72
Lead, oxidized	0.63
Cast iron	
Oxidized	0.62
Bright	0.22
Brass	
Oxidized	0.60
Polished	0.10
Nickel, oxidized	0.42
Zinc, oxidized	0.11
Silver, polished	0.03
Aluminum	
Polished	0.04
Oxidized	0.11
Paper	0.93
Aluminum paint	0.30
White enamel	0.64
Black lacquer	0.87

where h_b = watts per square inch, t = temperature gradient. These equations apply to vertical planes more than 12 in high. For heights less than 12 in, multiply the coefficient in Eq. (21-7) or (21-8) by the constant noted below.

Height, in	Constant
8	1.35
6	1.53
4	1.76
2	2.70

The rate of heat transfer by natural convection from a horizontal surface facing upward is about 35% greater and a rate from a horizontal surface facing downward is about 30% less than from a vertical surface. Hence, for average conditions, and on the sides of the body exposed, the average rate of heat transfer by natural convection is well represented by Eqs. (21-7) and (21-8).

If the natural-convection heat transfer is to a surface, the portion affected is the reverse of that stated for transfer from a surface. This does not affect the average rate of heat transfer.

Forced Convection. The velocity of the gas is a dominant factor in the transfer of heat by forced convection. The flexibility of its control, together with the penetrating property of gases, makes forced convection an effective method of heat transfer in many cases. Forced convection can also be applied to fluids.

In electric-heating practice the main limitation on the use of forced convection on gases is the upper limit of operating temperature of fans, about 650°C (1202°F). In some cases the gas can be heated after leaving the fan, and thereby higher temperatures than that noted above can be used.

There is in each use of forced convection a requirement of defined passages for the movement of the gas. This method also entails an investment charge for the fans, ducts, etc., and requires an expenditure of additional energy.

Some of the more common arrangements of surfaces with various conditions of forced convection have been studied, and relate to (*a*) gases flowing in pipes and (*b*) gases flowing past plane surfaces.

Finned Surface. If the heat-transfer coefficient of a surface is less than the heat-transfer coefficient which is the measure of the rate of flow of heat through the body, to or from the surface, the addition of fins or ribs to the surface will increase the rate of heat flow to or from the surface.

As a general rule, a spacing and dimensioning of fins that increase the surface area seven to nine times is good practice. The height of the fins should not exceed 3 to 4 times the spacing.

The principle of increasing the surface area per unit length of the surface by the addition of fins is equally applicable to the transfer of heat by radiation.

Heat Absorption. The rate of heat absorption for given external thermal conditions is determined by the diffusivity coefficient of the substance. This coefficient is defined as the ratio of the thermal conductivity to its heat-storage capacity per unit volume; thus

$$n = \frac{k}{p_c} \tag{21-9}$$

where k = thermal conductivity, p = density, and c = specific heat. This coefficient is but little affected by temperature.

The heat absorbed by a material for a given temperature rise is

$$q = \frac{w \times c_m \times t'}{3.412} + L + R \qquad \text{Wh} \tag{21-10}$$

where w = weight in pounds; t' = temperature rise in degrees Fahrenheit; c_m = mean specific heat over the temperature range; L = latent heat, if any, in watthours; and R = heat of reaction, if any, in watthours.

Heat of reaction here means the quantity of heat that must be added to the charge or subtracted from the charge, depending on whether the reaction is endothermic or exothermic.

The total heat required in a given case is the sum of the heat absorbed by the material, that is, the charge and the heat loss of the operation. The corresponding average power input is the total heat in watthours divided by the time of the heat cycle in hours. The power-time characteristic is individual for each class of heating surface.

Thermal Efficiencies. The conversion efficiency of the heating device is the ratio of the heat absorbed by the material, that is, the charge for a given temperature rise and the corresponding heat input to the device. If the charge is under cover, the heat absorbed by the material of the enclosure (the heating chamber) must be taken into account, along with the heat dissipated from the outer surface of the chamber. With increase in length of heat cycle the heat stored in the materials of the enclosure affects the conversion efficiency less and less, and with continuous operation it is usually negligible.

The operating efficiency expresses the combination of the conversion efficiency and the efficiency of the method of operation. If containers or removable supports for the charge are used, it is necessary to distinguish between net and gross operating efficiencies; the latter includes the heat absorbed by the containers or supports. A much-used rule of thumb for preliminary estimates is an operating efficiency of 50% for many heat applications under cover.

Thermal efficiencies are expressed as watthours per pound, pounds per kilowatthour, kilowatt hours per ton, etc., according to the scale of the operation.

High-Density Electromagnetic Force. Usually utilized to achieve fast heating times in air, vacuum, or gas atmospheres, this method of heating is more commonly known as induction heating. The electromagnetic force is developed at some frequency through the utilization of a suitable conducting coil. The coil is generally made of copper, often it is liquid-cooled, and it is designed to encircle or to be in close proximity to the material (nominally metal) to be heated. The high-density electromagnetic force coupled to the metal load induces current to flow in the load, and this current directly heats the material.

Concentrated Electrostatic Fields. Usually applied to fast heating times in air, vacuum, or gas atmospheres, this method of heating is more commonly known as dielectric heating. The electrostatic field is developed at a very high frequency. Through the utilization of metallic plates (often copper) placed above and below or on either side of the material to be heated (nominally a nonconductor) the electrostatic field causes a rapid reorientation of the molecules in the load, and heat is thereby produced internally.

Industrial Heating. This is classified as follows:

1. The heating of materials to temperatures below the temperature at which a change of state occurs
2. The heating of materials to cause a change of state, that is, evaporation and melting
3. The heating of materials to bring about a chemical reaction

21.12.2 Methods of Electric Heating

The methods of electric heating are

1. Resistance heating = I^2R effect
2. Electric arc
3. Induction heating = special application of I^2R
4. Dielectric heating

The resistance method and the induction method are widely employed. The use of the electric arc is generally confined to high-temperature applications. Dielectric equipment is widely used for the heating and welding of nonconductors.

Alternating current is used in most applications of electric heating. Direct current is required for a thermal process in which electrolysis is involved.

Resistance heating combines the basic attributes of electric heat with the ability to provide efficient temperature control up to 5425°F (3000°C). Apart from mechanical automation for material-handling systems, the trend to automate overall processing will continue.

I^2R effect is expressed by the equation

$$q = i^2rt = \frac{e^2}{r}t \qquad \text{Wh} \tag{21-11}$$

where r = resistance of the path of the current in the load; i = the current in amperes, for varying current, the rms value; e = applied voltage, across load; and t = time, in hours, of the flow in current.

The resistivity-temperature coefficient of metals is positive; that of nonmetals, negative. An exception among nonmetals is graphite.

The relation between applied voltage and change of resistance—a general case in heat applications—for constant power is

$$e_2 = e_1 \frac{z_2}{z_1}\sqrt{\frac{r_1}{r_2}} \tag{21-12}$$

or if the power factor of the circuit is unity:

$$e_2 = e_1 \sqrt{\frac{r_2}{r_1}} \tag{21-13}$$

where e_1 = initial voltage, r_1 = initial resistance, z_1 = initial impedance, r_2 = new resistance, z_2 = new impedance, and e_2 = new voltage.

The rate of heat development, that is, watts, is obtained by dividing the quantity obtained by Eq. (21-11) by the time period in hours or frequencies of an hour.

In indirect resistance heating, the heat is developed in resistors (heating units), a circuit apart from the charge, and is transferred from the surface of the resistors to the surface of the charge by one or more of the modes of heat transfer, namely, conduction, radiation, and convection (natural or forced).

If the heat transfer is by conduction, the resistor must be in contact with the charge. An enclosure for the charge (heating chamber) is required for heat transfer by radiation and convection.

21.12.3 Electric Heating Equipment

Resistors. The materials in general use for resistors for medium- and low-temperature surfaces are the two alloys B82 and B83 designated by the ASTM.

B82 Alloy. Additional data on round wires and rectangular sections are given in Table 21-7. Approximate maximum operating temperature is 1160°C (2100°F).

B83 Alloy. See the footnotes to Table 21-7. Approximate maximum operating temperature is 900°C (1652°F).

The melting ranges of these alloys are narrow. They do not soften if the temperature of the resistor is kept at a reasonable margin below the melting range. The coefficients of expansion are low, and the change of resistance with the change of temperature is small. The resistivities permit the use of reasonable cross sections and lengths at standard low-distribution voltages.

These alloys have the prime requisite of resistance to oxidation and scaling when operated in the air. However, this protection is not perfect, and it decreases in value with increases of resistor temperature. The alloys are susceptible to harm from compounds of sulfur and are affected to some extent by carbon monoxide. Unexpected causes of chemical action on resistors of these alloys are foreign gases or gases evolved from the material being heated.

TABLE 21-7 B82 Alloy (Ni-Cr) Average Values

No. B & S	Diameter, in	Surface area, in^2/lin ft	Ω/ft straight round wire, at 20°C*	Ft/lb bare round wire†
1	0.289	10.91	0.00778	4.18
3	0.229	8.65	0.0124	6.66
5	0.182	6.86	0.0196	10.5
7	0.144	5.44	0.0314	16.8
9	0.114	4.31	0.0500	26.9
11	0.091	3.42	0.0785	42.2
13	0.072	2.71	0.1250	67.4
15	0.057	2.15	0.200	107
17	0.045	1.71	0.321	172
19	0.036	1.35	0.502	26
21	0.0285	1.073	0.800	430
23	0.0226	0.851	1.27	684
25	0.0179	0.675	2.03	1,090
27	0.0142	0.535	3.22	1,730
29	0.0113	0.424	5.09	2,730
31	0.0089	0.337	8.21	4,420
33	0.0071	0.267	12.90	6,930
35	0.0056	0.212	20.70	11,150
37	0.0045	0.168	32.10	17,200
39	0.0035	0.133	53.10	28,400

*For B83 alloy, multiply by 1.037.
†For B83 alloy, multiply by 1.02.

In addition to the B83 alloy, several other similar alloys are listed by manufacturers for low-temperature service.

It is often desirable for low-temperature heating applications to use resistors known as metal-sheath heaters. These devices provide a metal sheath, usually a nickel alloy or a stainless steel, as a protective cover over the resistance device and insulated from it by a magnesia oxide. Sheath heaters are ideal for direct immersion and conductive heating, as the sheath tends to assume the temperature of the resistance element. Direct contamination of the resistance element is thus avoided.

Resistors for operating temperatures above 1150°C (2100°F) are made of silicon carbide, molybdenum, tungsten, and graphite.

Silicon carbide is the basis of a resistor material for operating in air for temperatures up to about 1500°C (2732°F). The material is formed into rods of diameters and lengths for combination into circuits of required electrical rating. The range between the cold and hot resistances is small enough to permit single-voltage operation. The resistance gradually increases with use, and an autotransformer with taps to compensate for this increase of resistance is desirable.

Molybdenum resistors are suitable for temperature up to 1650°C (3002°F). This metal is ductile enough at room temperature for drawing into wire for resistor windings. The wide range between the cold and hot resistances makes multiple-voltage operation necessary. The supports (insulators) of the windings should be of magnesia or zirconia. Molybdenum resistors cannot be operated in air and also must be protected against reactions with silicon and carbon. The metal is immune from reactions with sulfur, nitrogen, hydrogen, and water vapor. A hydrogen atmosphere is ordinarily used for the protection of these resistors. Molybdenum is not suitable for resistors of vacuum furnaces because of its high vapor pressure.

Tungsten resistors can be used for temperatures up to 2000°C (3632°F). The maximum temperature is limited by the refractory supports of the resistor. The metal must be heated for drawing into shapes. The resistivity-temperature characteristic is practically the same as that of molybdenum. The low vapor pressure of tungsten makes it useful for resistors of vacuum furnaces.

Graphite resistors are suitable for any temperature that can be used. The resistors must be protected against oxidation above about 600°C (1112°F). Also because of the chemical activity of carbon, special consideration must be given to the atmospheres surrounding the resistors. The material is available in a wide range of shapes and dimensions.

The *useful life* of a resistor (barring accidents) is dependent mainly on the temperature at which it is operated. That temperature must be high because of the required rate of heat transfer per unit area of the surface. Thus the operating temperature of a resistor is a function of the surface power density (watts per unit area) of the resistor. For a resistor of given shape and dimensions the operating temperature will vary with the watts input, since the surface area remains the same. Hence, watts per unit area is a first consideration in resistor design. This value in practice varies according to the conditions of use.

A rectangular shape for the resistor section gives a greater surface area for a given resistance than the round shape (Figs. 21-11 and 21-12); hence, the frequent use of the ribbon, or strip, form of resistor, particularly for the upper range of temperature. The cast resistor for furnaces also embodies this principle.

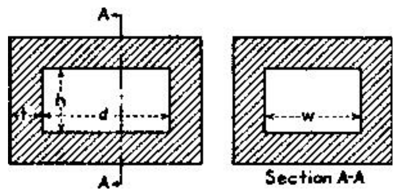


FIGURE 21-11 Shape factor—rectangular enclosures.

The *maximum voltage* of resistor circuits is limited by electrical insulation and elevated temperatures and by safety considerations to 600 V. This value may be exceeded in special cases. Standard distribution voltages and standard frequencies are used. Resistor windings may be in one, two, or more circuits and may be connected either single-phase or polyphase.

As a general rule, the *heat-equalizing effect* of the heat-storage capacity of the charge into the surrounding enclosure (if any) makes permissible voltage fluctuations within reasonable limits. However, the average voltage must be equal to the normal voltage to maintain the normal heating rate. The relation between heating rate and voltage is shown in Fig. 21-13.

The *power input* to resistors supplied from constant-voltage circuits can be varied by the use of multiple circuits and series-multiple-switching combinations. Another method is the use of an autotransformer or transformer with taps. With 3-phase resistor circuits the change from delta connection to wye connection, and vice versa, is available. In each case the power input to the resistor is proportional to the square of the voltage. For example, with the change from delta to wye the power input is reduced to one-third its value with the delta connection.

Other devices are available for power controls such as saturable reactors, regulated or unregulated; variable inductances; and stepless variable autotransformers.

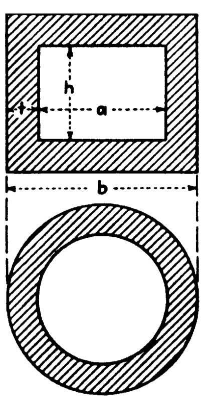


FIGURE 21-12 Shape factor—cylindrical enclosures.

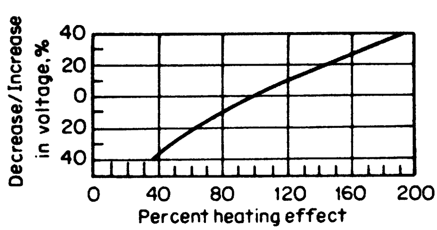


FIGURE 21-13 Variation of heating effect with changes in voltage.

Self-Contained Units, Water Heating, and Steam Generators

Resistor heating equipment includes many small devices such as electrically heated tools, appliances of various kinds, self-contained units, ranges, water heaters, stills, sterilizers, and baths. The larger units include ovens, kilns, lehrs, furnaces, etc. Each of these descriptive terms relates to a field—some more general than others—of electric heating service.

Small devices are often used. As a rule the best design of a resistor for any small device is obtainable only by trial. The space is generally limited, and the conditions which affect heat transfer vary widely. For many devices a useful guide is 20 W/in^2 of resistor surface. Both higher and lower values are used according to the conditions in each case. Metal-sheath heaters can be obtained for operation at many wattages to be operated at specified voltages of 115 or 230 V.

Self-contained units meet manifold needs for localized heat in small quantities within the lower temperature range and extending up to about 815°C (1500°F).

The construction of these units consists of a resistor of the 80 Ni–20 Cr alloy embedded in a heat-resisting material such as powdered fused magnesia and enclosed in a cylindrical metal sheath. In many cases they are applied for direct-contact heating, that is, for heat transfer by conduction.

The forms of the self-contained heating units include insertion (or cartridge) units, straight or curved tubular units in a shape to conform to the method of use, space (or strip) heaters, immersion units for heating liquids, fin-type units for air heating, cast-in units, and a special construction for melting soft metals.

Standard voltages of self-contained units are 115, 230, and 460 V. Space restrictions of the terminals do not permit higher voltages except for special designs. The ratings of these units range from 30 W to 10 kW.

Space Heaters. These are made in two types:

1. With steel sheath; specific rating 10 W/in^2 of surface; maximum operation temperature 399°C (750°F)
2. With porcelain-enameled steel or alloy-steel sheath; specific rating 15 W/in^2 surface; maximum temperature 650°C (1200°F)

Both types are used for ovens, warming tables, process machinery, space heating, etc.

The rating in watts per square inch of affected area of the sheath of the unit is considered in relation to its application: (1) the resistance to oxidation or other chemical action and (2) the rate of heat absorption by the surrounding medium. Immersion units for heating water have comparative ratings, 40 to 50 W/in^2. However, these units should always be installed so as to permit free circulation of water, and they should not be operated unless completely submerged.

Oils in general do not absorb heat as rapidly as water. Also, immersion units in oil accumulate carbon deposits on the sheath. Hence, units for this service have a much lower rating—20 to 25 W/in^2. Steel is a satisfactory sheath material for mineral oils. Other oils must be given individual consideration with reference to the chemical action on the sheath material. The same consideration applies to heating chemical materials in liquid form; for example, a unit with a lead sheath is used for nickel and copper solutions.

The temperature-watts characteristic of the fin-type units is given in Fig. 21-14.

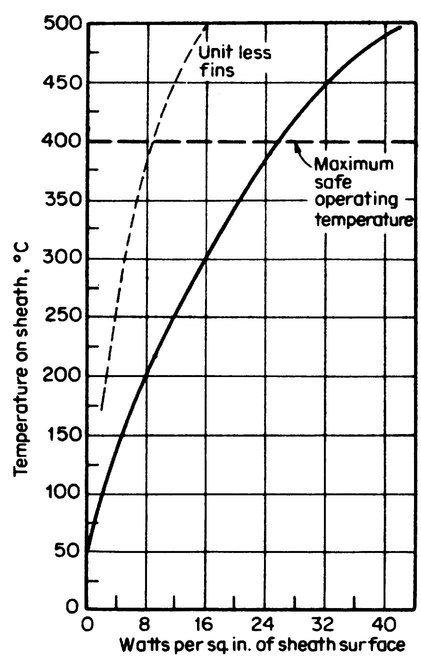


FIGURE 21-14 Temperature-watts characteristics of fin-type unit in still air.

Room Heating. The space heater is well adapted to heating rooms; radiation and natural convection are used for the heat transfer.

An approximate equation for the rate of heat loss from structures or ordinary building construction is

$$\text{kW} = \frac{(0.02NV) + (1.13G) + kA}{3412}(T_r - T_0) \tag{21-14}$$

where N = number of air changes per hour (usually two or three)
V = room volume, ft^3
G = area of windows, ft^3
A = area of ceiling and exposed walls, ft^2
$T_r - T_0$ = temperature difference, °F, between inside and outside
k = heat-transfer coefficient, Btu/h · ft^2 · °F

Average Values of k

Wood construction	0.25
8-in brick wall	0.50
12-in brick wall	0.36
6-in concrete wall	0.79
10-in concrete wall	0.62

The heat-transfer coefficient can be reduced by the addition of proper insulating material in the walls and ceiling of the room. Other heat sources must be considered in determining the actual heating-equipment requirements, such as lighting, appliances, machinery, and people.

A heating cable with a lead sheath for temperatures up to 74°C (165°F) has a variety of uses, for example, warming soil to promote plant growth, prevention of freezing of liquids in pipes, and keeping gutters and downspouts open.

Electric Steam Generators. These are made in three types:

1. With immersion heating units up to 300 kW, single-phase, standard low voltages
2. With 80 Ni–20 Co alloy resistors mounted in the tubes of a fire-tube boiler; sizes up to 1000 kW, single-phase or 3-phase, standard low voltages
3. Electrode type: large units 200 kW and above; all 3-phase, either 2300 or 6600 V

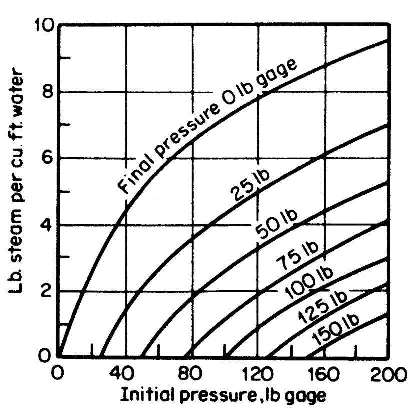


FIGURE 21-15 Flashing capacity of steam generators.

The heat loss from a generator in operation is the sum of the loss through the heat insulation and the loss by conduction through the pipe connections. The time of heating is very short and the total loss within a given time is small compared with the heat output in the steam. Hence, the insulation of the shell can be moderate in amount for a high conversion efficiency—95% to 97% and higher. This efficiency is usually given as 3 lb of steam/kWh. The boiler-horsepower rating is one-tenth the kilowatt rating.

There is little storage capacity in a steam generator, and the power demand cannot be reduced by storage as in water heaters. When a quantity of steam beyond the rating of the generators is required, it cannot be obtained by releasing the steam through a pressure-reducing valve. The amount of steam thus available per cubic foot of water in a generator can be read from Fig. 21-15.

The auxiliary equipment of electric steam generators is the same as that of fuel-fired steam generators. However, full automatic control is more frequently used with the electric steam generator. The main use of the small sizes of electric steam generators is for the supply of steam at points which cannot conveniently be reached by steam pipelines and in locations where the fuel-fired unit is not permissible.

The larger units have been applied mainly for utilization of the excess water capacity of hydroelectric plants and by industrial plants which require large quantities of processed steam, for example, paper mills.

Resistor Ovens

The two types of resistor ovens are convection and radiation.

An *oven* is a low-temperature heating chamber with provision for ventilation.

Oven service is drying and baking. Drying is the removal of liquids by evaporation. The rate of drying is limited by the rate at which the liquid will move from the interior of the mass. This is a limitation on the power input. The effect of a too high power value is skin drying.

Baking is a hardening process which involves chemical and physical changes in the material. If a material contains a solvent, much of that component is removed during the first stage of baking. Air drying preliminary to baking may be advantageous.

The technique of baking is a correlation of temperature and time. The general rule is: The higher the temperature, the shorter must be the heat cycle. The temperature-time relation for each material is individual and must be determined by trial.

Ventilation. As a rule, the ventilation of oven chambers is forced by means of fans. The removal of inflammable vapors is a special procedure.

Convection, Box, and Continuous (Air) Ovens

Convection Ovens. The transfer of heat is by forced convection in gases, mainly air. Steel panels are used for the enclosure of the heating chamber. A double wall with inner heat insulation is general practice. The heat-storage capacity of these enclosures is small.

Box Ovens. Box ovens—batch or stationary heating—are of two designs:

1. Heating chambers with hooks, racks, or shelves for the support of the charge
2. Heating chambers arranged to receive drawers, trucks, or similar removable supports for the charge

Continuous (Air) Ovens. These ovens are flexible in design. The structure may be horizontal, with or without inclined entrance and exit, or vertical. The travel of the charge may be one-way or a U-shape movement. The latter permits the absorption of heat from the outgoing material by the incoming material. The chain conveyer is applicable generally.

Radiation Ovens

In radiation ovens, heat is transferred by radiation. The heating units are tungsten-filament lamps (heating lamps) with self-contained or external reflectors.

Standard heat lamps are 125, 250, 375, 500, and 1000 W, 115 V. The self-contained reflector lamps of the first three ratings are most generally used. The filament temperature around 2500 is low enough to give a comparatively long life, average 5000 h or more. The large part of the radiation from filament lamps is in the infrared region of wavelengths; hence the term *infrared heating.*

Construction of the radiation oven is similar to that of the convection oven. The main difference is the detail incident to the lamp-type heating units. The general arrangement is a bank of heat lamps which surrounds or partially surrounds the charge. In a continuous oven this bank forms a tunnel through which the charge travels. This type of oven is frequently built with metal-sheath heaters as a source of infrared energy. High temperatures are available through the use of quartz-infrared units, used particularly in higher wattage ratings.

General service is drying and baking processes which require temperatures below 315°C (600°F). The method is best adapted to continuous horizontal ovens for heating charges, which present a large surface area in proportion to the mass, for example, sheet metal, textiles, paper, and some plastics, and require only surface heating.

Absorption coefficients of the surface of charges should be reasonably high. Correspondingly, the inner wall surface of the heating chamber should have a low absorption coefficient. Aluminum linings are used in many cases.

Concentration of radiation (power density) on the charge surface depends on the wattage and spacing of the lamps; the efficiency of the lamp reflector units; and, to a limited extent (within the range of practice), the distance between the lamps and the surface of the charge.

21.13 ELECTROMAGNETIC INDUCTION

Classification. Electromagnetic induction is a process used to heat metals for purposes of hot working, heat treatment, welding, and melting.

In the hot-working area, induction is used largely for through preheating prior to forging and extrusion; it has been found to be economically attractive as an alternative to conventional gas-furnace heating processes. The application of induction for slab preheating, although in commercial use, has not attained as high a level of acceptance.

The prime application of induction in the heat-treating field is in the area of surface hardening. It is widely accepted as the principal means of surface hardening for a wide variety of mass-produced parts. In some applications, competition appears to be coming from ion nitriding, laser hardening, electron-beam hardening, and high-frequency resistance hardening. By and large, the first three of these processes do not appear to offer a threat to the induction market. The development of high-intensity induction surface hardening and high-frequency resistance heating and hardening (which employs similar radio-frequency equipment) can be employed to produce the very shallow case-hardened depths achievable by the other three processes.

The use of induction for through heat treatment is less widespread than for surface hardening. This is particularly true for some processes as tempering and annealing. Recent research work has suggested that such applications may indeed be feasible with induction, however.

High-frequency induction and resistance welding are considered to be well-established processes. High production rates and low cost are two of the main factors which make it more attractive than competitive processes such as arc welding in high-volume areas such as pipe and tube production.

Although induction heating is a well-established technology, development efforts in such areas as higher-efficiency equipment, coil design, and automation should enable the process to retain and possibly to enlarge its markets.

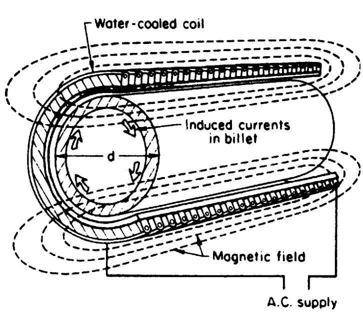


FIGURE 21-16 Basic principle of induction heating illustrated for a tubular conductor heated by a solenoid coil.

Principle of Operation. The material to be heated is exposed to an alternating (or varying) magnetic flux. The material to be heated must be a conductor so that currents are thus induced in the material (or load) and flow wholly within its mass. Generally this type of heating is produced by an electromagnetic field caused by an alternating-current source. It is possible, however, to produce similar results, but at lower current densities, with a high-flux permanent magnet moved at high speed across conductive materials. The term *eddy-current heating* is sometimes applied to the method.

Assemblies. One of the simplest induction-heating setups is shown schematically in Fig. 21-16. Here, the induction coil, or simply inductor, is a helical (or solenoidal) copper conductor carrying an alternating current which surrounds a tubular workpiece. Associated with the ac is a magnetic field whose flux lines pass through the workpiece. The strength of the magnetic field varies with the ac; in so doing, eddy currents are induced in the workpiece. Generally, the induced current path is parallel to that of the currents in the inductor.

For a given workpiece and coil arrangement, the heating rate can be adjusted by increasing or decreasing the induced current in the workpiece. This is done by adjusting the coil current which is controlled by the ac or induction generator. If the workpiece is magnetic (for example, steels at low temperature), additional heat is produced by hysteresis losses. Such heating is usually small but, in some cases involving strong magnetic fields, it can become significant.

Shape of Load. The ideal load shape is a cylinder. It may be a solid, or a tube. Shape alone does not affect the principle involved. It may introduce some complications in geometry. Any shape load can be heated by the inductive method, but not always to obtain the desired results. Special techniques are required for some unusual shapes.

Basic Relations. For a cylindrical load and other loads where the coil encompasses the load, the directions of the magnetic flux are parallel to the longitudinal axis of the load and hence to its lateral surface. The factors which, for a given intensity of longitudinal magnetic flux at the lateral surface of the load, determine the rate of heat development in the load are

1. The electric and magnetic properties of the material
2. The frequency
3. The radius, or one-half the thickness of the charge

Heat Concentration. In considering a solid load, it is necessary to visualize the current subject to skin effect induced to flow in circular paths parallel to the circumference of the load. Each circuit possesses inductance, and there is a subsequent progressive decrease in current strength radially. The rate of decrease changes with frequency.

The Rate of Heat Development. The general equation for the rate of heat development by symmetrical currents in a long load (length much greater than its diameter) of any shape for a longitudinal and sinusoidal alternating magnetic flux at its lateral surface is

$$P = H^2SL \times \sqrt{\mu f} \times \sqrt{p_a \times 10^9} \times 10^{-7} \qquad \text{W} \tag{21-18}$$

$$H = \frac{4\pi N \times \sqrt{2}i}{10L} \qquad \text{Oe} \tag{21-19}$$

where p_a = resistivity of the material of the load in ohm-centimeters cubed
μ = permeability of material of the load
S = shape factor of the load
L = length of the load in centimeters
N = number of turns in the coil around the load
i = current in the coil, A

Magnetic Charges. The ac excitation curve of the magnetic material more or less coincides with the dc magnetization curve as saturation is approached. Hence, the latter curve (usually more readily available) can be used as a reference in induction heating. For values of H higher than that required for saturation, the value of μ corresponds to the saturation values of p of H.

Permeability. Permeability is dependent on frequency up to around 10 kHz. Above that frequency, permeability decreases to some extent with increase of frequency, but this effect is not well defined.

Hysteresis. In induction-heating practice, the rate of heat development by hysteresis in a magnetic material at room temperature is a small percentage of the eddy-current watts. The hysteresis watts decrease rapidly with rise of temperature. Hence, hysteresis is a negligible factor in the inductive method of heating.

Primary Coils. Single-layer coils are standard practice. Tubular conductors—round, oval, or rectangular—provide for water cooling of the coil.

The number of turns is selected with reference to the desired relation of amperes and volts for a given service. The restrictions on this choice are the space required for conductor insulation and the requirement of water cooling.

21.14 ELECTRIC WELDING

21.14.1 Resistance Welding

Resistance welding includes spot welding, projection welding, seam welding, and butt welding. All are based alike on the principle of resistance heating but differ in the details of application.

Spot welding is applied to overlapping sheets. These are clamped between water-cooled electrodes, pressure is applied, and a current impulse is passed through the assembly. The material in the zone of pressure is heated to fusion, and the joint thus made to cool under pressure.

Spot welding is largely employed on sheets 0.025 to 0.125 in thick, although welds on thicker sheets and plates up to 1 in thick are made regularly.

The area of the spot weld is small; weld diameters of 0.25 to 0.625 in are the usual range. The maximum diameter is limited by the area of uniform contact between sheets that can be obtained with practical pressure.

Rapid heating by using high current values is necessary to bring the metal that is to form the joint to the required temperature without more than normal heating of the adjacent metal. Current values range from 5000 A upward. The voltage between the electrodes is usually less than 2 V. The open-circuit voltage is generally less than 12 V. The time period of current flow varies widely, depending on the thickness of the sheets, kinds of material, etc. For thin steel sheets this period is about 1 c ($^1/_{60}$ s) for each 0.010 in of the total thickness of the two sheets to be joined. The pressure, current value, and time are variables that must be correlated in each case.

Spot welding with a single current impulse is limited to sheets not thicker than about 0.125 in. Because of the heating of the electrodes, and consequent short life, with heavier material, thick sheets are welded by using a number of current impulses, for example, 4 c on, 2 c off, repeated as required. The no-current intervals, that is, cycles off, prevent the overheating of the electrodes.

Projection welding differs from spot welding only in the use of the buttonlike projection on one of the sheets to define the path of the current. A number of adjacent welds can be made simultaneously without interference, a useful-method for assembly work.

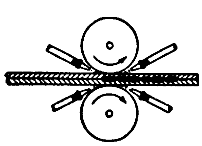


FIGURE 21-17 Seam-welding electrodes with water cooling.

Seam welding can be made by putting a series of spot welds on thin sheets by passing the sheets between two pressure electrodes (Fig. 21-17) and with continuously interrupted current. Speeds range from 10 to 400 in/min; approximately 60 in/min is the most common speed. Twelve welds per inch with 6- to 10-in-diameter electrodes is representative practice.

Sometimes classed as seam welding but otherwise known as edge welding is the type of welding accomplished on continuous tube mills, or continuous butt-welding installations where the edges of metal are brought together and welded as a seam. In this type of welding pulsed currents are not used to obtain the weld, but rather a continuous ac or dc supply fed through rotating electrodes to the abutting edges of the seam produces the electric energy required for the weld.

The *variables* in these three methods of welding which must be correlated are

1. Electrode diameter, material of electrodes, contact area, cooling, and pressure
2. Nature and surface of metal or metals to be welded
3. Timing of current flow
4. Current value and waveform

Timing circuits for resistance welding are either (1) nonsynchronous or (2) synchronous (relating to the welding current zero).

Nonsynchronous timers are of various types, ranging from the simple application of foot pressure to close and open the welding circuit, according to the judgment and skill of the operator, to the combinations of cams, limit switches, etc., the functions of which are independent of the operator.

The synchronous timers are electronic, using thyratrons and ignitions. The most simple of these are those which are adjustable, as regards to current impulse, in 1-Hz steps. When used with an ac supply, the ignitron contractor will pass current each half cycle that the anode is positive. A second ignitron is connected inversely to pass current during each of the other half cycles.

Electronic control for resistance-welding circuits is inherently a precision device. The choice between the two types of control is primarily a matter of the value of exactness in timing, which becomes more important as the welding time is decreased.

Current control in welding circuits, that is, the control of the rate of heat development, is obtained by taps in the primary winding of the welding transformer or by phase control of the welding current, retarding the point at which current starts to flow in each half cycle beyond the power-factor angle. A combination of the two methods is often used.

Percussion welding is a self-timing spot-welding method. A current impulse is obtained by the discharge from a capacitor or from a magnetic field. (This method is usually classified under dc resistance welding.) The capacitor arrangement for percussion welding is used extensively in the manufacture of lamp filaments.

Butt-Welding Upset Method. End-to-end welds, lap welds, and butt welds are included here. The faces of the parts should be prepared for even contact. The parts are clamped and brought together by hydraulic pressure, thus closing the circuit. Plasticity occurs simultaneously over the entire contact area. The joint is cooled under pressure. The method has a wide variety of applications, large and small. In continuous pipe welding, the current is conducted to the work by a pair of rotating electrodes. This, as previously mentioned, is sometimes referred to as the seam-welding technique. Butt welding can be obtained by utilizing induction-heating equipment; how this can be accomplished will be described later.

Flash Welding. No preparation on the face is necessary. One face is stationary, the other movable. Voltage is applied through pressure contacts. A momentary contact establishes current flow and flashing (the melting and vaporization of the minute contact points). The movable face is fed forward continuously at the rate at which the metal is melted. The pressure of the gases developed between faces prevents the access of air to the hot faces and blows out the molten metal. The metal adjacent to the faces becomes plastic, and the joint is made by fusing the parts together by hydraulic pressure. The circuit is opened by a jam relay as the faces make contact. However, if a considerable amount of flashing is required to remove impurities, the circuit is held closed for a predetermined time and may be opened just prior to the application of the last pressure value.

21.14.2 Arc Welding

Welded joint is a union of metal parts made by localized heating, without pressure in arc welding and supplemented by pressure in resistance welding. The technique of welding is built around the methods of concentration of heat and is based on the metallurgy of welds in the mechanics of joints.

Arc-Welding Manual Operation. Methods of arc welding are classified as follows:

1. Direct-current arcs, one electrode—metal or carbon

  a. Unshielded arc
  b. Shielded arc

2. Alternating-current arcs, metal electrodes

  a. Unshielded arc, one electrode
  b. Shielded arc, one electrode
  c. Shielded arc, two electrodes (atomic-hydrogen process)

Joints. The five principal types of arc-welded joints are butt joints, corner joints, edge joints, lapped joints, and T joints. Any length of joint can be made by travel of the electrode; one or more passes may be required.

The position of the joint (see Fig. 21-18), especially for multipass work, is an important factor in the economics of welding. The maximum welding speeds obtained with the flat position are about twice as fast as those with the horizontal position and about 4 times faster than the maximum speeds obtained with the overhead and vertical positions. Welding in the flat and horizontal positions is much easier for the operator than overhead welding—so much so that the arc can be kept in operation about twice as long per hour of overall time. Likewise, vertical welding is less tiresome than overhead welding and represents an increase of about 50% in welding time without an increase in the fatigue of the operator.

The *weldability* of metals and alloys can be stated as easy, intermediate, and difficult to weld. This term, when used with reference to carbon steel, relates to carbon content. Carbon steels up to about 0.30% C are easy to weld. That class of steels is used in a large part of the fabrication of materials by welding. Practically all metals and alloys can be electrically welded, but in many cases an individual study of the metal or alloy, dimensions, and position of the weld is required to develop the particular technique necessary.

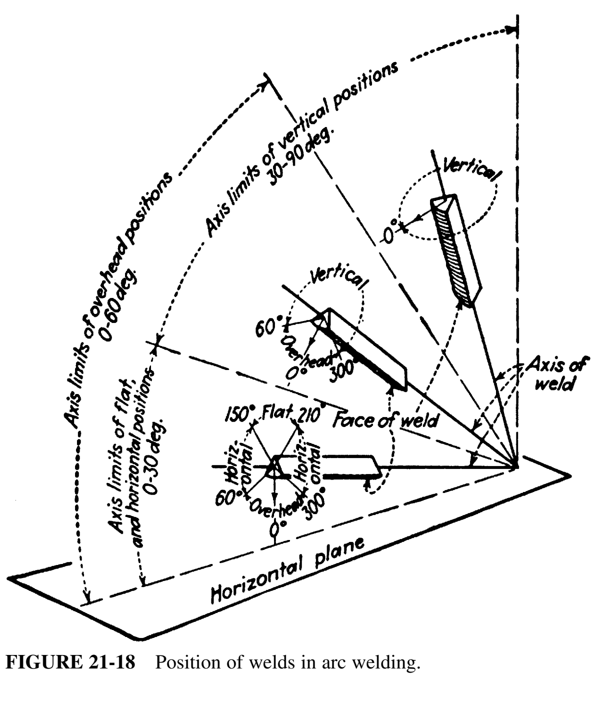


FIGURE 21-18 Position of welds in arc welding.

A characteristic feature of arc welding is that symmetrical sections (I beams and T beams) can be used.

Single-Electrode Welding. The arc is maintained between the electrode and the work. Round steel electrodes, with carbon content suited to the class of work, supply the metal (filler metal) for the joint.

The phenomenon of the transport of the molten metal from the electrode to the work is dependent on the position of the joint. It is, however, a factor in the selection of the electrode for each joint position. The making of a sound weld requires that the filler metal be deposited only on molten-base metal. The depth of the molten-base metal is termed *penetration.*

Metal Electrode Types. There are two basic types:

1. Bare electrodes (or lightly coated electrodes)

2. Coated electrodes

The bare electrode is covered with a film or its oxide to aid in stabilizing the arc circuit. This type of electrode is used mainly for joints which carry only static loads; the coated electrode has a jacket made of either organic (combustible) or inorganic material or a combination of the two types. This coating serves several purposes:

1. The organic material provides the gaseous shield to restrict the access of oxygen and nitrogen (air) to the molten metal.

2. The inorganic material makes a slag to refine the molten metal. (The principle is the same as the refining of steel and the manufacture of steel in an arc furnace.) This slag also serves as an insulating blanket to retard the rate of cooling of the weld metal.

3. The gas envelope and the slag aid materially in the stabilization of the arc circuit.

4. In some cases the coating is used to add alloying elements to the weld metal.

Electrode coatings are of different types, corresponding to the class of weld service (kind of base metal, joint position, etc.). The selection of electrode size, carbon content, and type of coating is a major consideration in welding practice.

Powdered-Metal Electrodes. Some electrodes are available which have powdered iron, which will make up a good portion of the metal to be deposited in the joint, mixed in with the flux coating. This makes the electrodes much larger on the inside for the same-diameter core wire as standard flux-coated electrodes. These electrodes are well suited to down-hand welding and are said to deposit more metal in a given period of time than standard flux-covered electrodes.

Low-Hydrogen-Type Electrodes. Some types of steel welding rods have special treatment during manufacture to reduce the amount of hydrogen in the coating. The reduction of hydrogen in the coating produces a weld with little or no absorbed hydrogen in the base metal and reduces the likelihood of underbead cracking.

Electrode Quantities. The weight of electrode metal required for a given joint is

$$W = \frac{\text{weight of metal deposited}}{l - \text{electrode loss}} \tag{21-20}$$

The weight of metal deposited is calculated from the volume required to fill the joint plus reinforcement, if any. The electrode loss is the sum of the scrap end loss, the spatter loss, and the coating loss. The scrap end loss varies with the diameter and length of the electrode. With 14-in electrodes the loss varies from 10% to 20%. The average value of this loss is about 17%.

The other losses depend on the type and size of the electrode, joint position, welding current, arc voltage, and skill of the operator. With bare electrodes these losses vary from 8% to 15%, average 13%; with coated electrodes, from 15% to 35%, average 27%.

Carbon (Graphitic Carbon) Electrodes. These are used with direct current for arc welding of nonferrous metals and alloys and with the addition of external shielding (gases from material placed adjacent to the arc) for welding steel. The weld metal is supplied by melting a portion of the base metal. If additional metal is required, it is supplied by a filler rod. Carbon electrodes are not suitable for vertical and overhead welding. The use of a carbon electrode is declining.

The DC Arc. The dc arc is adapted to practically any class of welding service. The arc voltage varies with the type of electrode, ranging from 15 to 40 V. Both straight polarity (work positive, electrode negative) and reverse polarity (work negative, electrode positive) are used, depending on the character of the work. (Polarity here relates to the difference in the rates of heat development at the anode and cathode of the dc arc.) Metal electrodes should be selected with reference to the polarity arrangement. Carbon electrodes can be used with either arrangement; straight polarity is more common practice.

The rotating dc-generator single-operator unit, with a drooping voltampere characteristic, driven by any type of prime mover, is made in sizes ranging from 159 to 600 A, usual load-voltage rating 40 V, all sizes. A wide range of adjustments of welding current is provided in each size. The larger part of dc welding is done with this type of equipment.

Multiple-operator equipment and an adjustable series resistor for each arc circuit for operation from constant distribution systems are used to some extent. The special field for this equipment is where the service of each arc-welding circuit is infrequently in use, say, 25% of the time or less. In such cases the loss in the series resistor is less than the no-load loss of the single-operator unit for the same service.

Another type of dc-welding power source is a dry-type rectifier used in conjunction with a multiphase high-leakage-reactance transformer. Many of these rectifier-type welders use selenium rectifiers which are forced-air-cooled.

Rectifier-type welders are said to combine some of the desirable arcing characteristics of dc welding, such as easy arc starting, with those of welding transformers, such as reduced no-load losses. The same range of rating is available in rectifier-type welders used in rotating dc welders.

Alternating-Current Arc Welding. The field of ac arc welding is practically the same as that of dc arc welding.

The standard equipment for single-operator ac arc welding is a single-phase standard frequency transformer, sometimes with a number of voltage taps, and an adjustable reactance. Generally, capacitors are included to correct the power factor of the circuit to about unity between one-half and three-fourths load. The welding leads are limited to about 300 ft radius because of voltage drop.

Standard equipment includes ratings with the range of 150 to 1000 A, 30 to 40 V.

Atomic-Hydrogen Arc Welding. This method uses an ac arc between the tungsten electrodes (not including the work) in an atmosphere of hydrogen (shielded arc). Either commercial hydrogen or disassociated ammonia can be used. The weld metal is supplied by the base metal or, if additional metal is required, by a filler rod.

The method is very flexible. Practically any metal or alloy, ferrous and nonferrous, can be welded. Its greatest usefulness is for the fusion welding of certain ferrous alloys, for example, chrome-nickel steels, aluminum, and duralumin, but it must use flux to weld some of these materials. The welding of thin sheets, the production of tubing, and the repairing of expensive tools and dies are some of the common uses of the method.

Inert-Gas-Shielded Arc Welding. The method comprises the use of a single tungsten electrode, filler rod, and argon or helium as a gas shield. The particular field for the method is the welding of the light metals. No flux is required. The technique varies with the type of work. The following procedures are typical.

Welding aluminum and aluminum alloys—alternating current and argon

Welding magnesium and magnesium alloys—either alternating current and argon or direct current, reverse polarity, and either argon or helium

Welding stainless steel, mild steel, copper, and copper alloys—direct current, electrode negative, and either argon or helium or alternating current, high frequency, stabilized with argon or helium.

Consumable-Electrode Gas-Shielded Welding. This type of welding uses small, bare electrodes in coil forms. The welding is done in gas atmospheres. The 0.020- to 0.125-in-diameter electrodes are fed mechanically into a handgun for manual welding at feed rates up to 1000 in/unit. The voltampere arc characteristics of this type of welding are rising. Standard dc welding generators do not give the best results. Special low-voltage dc welders give better results. The optimum characteristic is one which is rising, to match the arc characteristic.

Argon, argon and 1% oxygen, helium, and carbon dioxide or mixtures of these gases are used. Argon is used for welding aluminum, magnesium, copper alloys, and titanium. Argon-oxygen or carbon dioxide is used for mild-steel welding. Almost all welding is done with the electrode positive.

The automatic consumable-electrode gas-shielded equipment does not require elaborate control systems. With the low-voltage generators the current flow automatically regulates itself to burn off whatever amount of wire is being fed into the arc.

As a result of this type of welding the generator produces the current to make the burnoff of wire rate equal to the wire feed speed.

Direct versus Alternating Current. The considerations that affect the choice between direct current and alternating current are mainly the following:

1. Alternating current has a higher (85%) efficiency. Rotating direct current and rectifier direct current are about 65% under load. Also, no-load losses for rectifier direct current and alternating current are less than the no-load loss of the single-operator dc equipment. The saving is 25% to 30%. A larger saving may be made compared with multiple-operator dc equipment.
2. The power factor of alternating current and rectifier direct current is low (about 40% lagging) and can be corrected by the use of capacitors.
3. The magnetic deflection of the arc (magnetic blowout) is not so pronounced with the ac arc. This phenomenon is often troublesome in working with dc arcs of high current values. The ac arc permits higher current and larger electrodes in such cases, with a consequent higher welding speed.
4. The cost of electrodes must be compared.
5. The cost of equipment must also be compared.
6. Where the range of work in a shop is widely varied, the use of both types of arc welding may be found desirable.

Automatic Arc Welding. Automatic arc welding, either dc or ac, is confined mainly to simple joints which can be placed in the flat position for welding.

The electrode for direct current may be bare or coated. Coated electrodes for alternating current are usual practice except when the flux (in granular form) is applied to the work (submerged-arc method). Coated electrode wire in long lengths for feeding from a reel has a slot in the coating for electrical contact. Coated electrodes in short lengths with end contact are also used by adapting the feeding mechanism for automatic return to the starting position when the electrode is consumed.

The power-supply equipment, dc or ac as the case may be, for automatic arc welding is the same as for the corresponding manual operation.

The automatic regulator controls the rate of feeding of the electrode to match the rate of melting of the metal electrode by controlling the speed of the electrode motor in response to slight changes in the arc voltage so as to maintain a practically constant average arc voltage.

One type of automatic regulator in wide use is an electric device which contains the value and the direction of armature coil of a dc electrode motor. This control uses two thyratron tubes in the armature circuit of the electrode motor, passing current from an ac source in one direction or the other, as required, to feed the electrode up or down. The magnitude of the current passed is controlled by grid-phase shift of the thyratron tubes in response to changes in the arc voltage in comparison with the preset reference voltage.

Automatic arc welding, a method independent of the human element in the operation, has many merits where it is applicable. The most important of these are uniform welds, high production rate, and low electrode loss. Also, a new trend is toward the use of robot welders with solid-state circuitry and logic control that perform repetitive functions and handle heavy production parts.

21.14.3 Induction Welding

Induction Resistance Welding. This method of welding incorporates the use of induction heating principles along with the application of high-frequency energy, usually in a radio-frequency range. It is particularly useful as a method of welding in tube mills. Here energy can be induced into the slit-formed tubing with a single-turn inductor and current caused to flow across the gap between pressure rolls where resistance welding is caused by the I^2R drop across the joint between the weld rolls.

This method of welding has been and is being used effectively for the welding of carbon steels, aluminum, copper, titanium, and mechanical stainless steel. This technique has been applied to portable weld units where it is possible to weld tubing and pipe on site from skelp.

The radio frequencies normally used for this type of welding are in the neighborhood of 200 to 450 kHz. It is possible to weld heavier wall material in this manner, frequencies being used in the neighborhood of 10,000 or even 3000 Hz, although these frequencies will not give as efficient an operating installation as the radio-frequency installation.

High-Frequency Resistance Welding. The application of contacts to line pipe utilizing radio-frequency energy has been successfully accomplished, and a weld similar to that obtained with induction resistance welding is thus secured. In this type of welding installation, high-frequency, preferably radio-frequency, energy is passed through sliding contacts to the V-shaped edges of the tubing. This current is then caused to flow across the contact between the pressure rolls, giving the I^2R heating effect of resistance welding at this point.

This method of welding is also effective on nonferrous materials, as well as carbon steels. It has been adapted to structural shapes and the finned-tube shapes.

Induction Butt Welding. Two properly faced tubular sections or rectangular sections butted together can be induction-heated with a suitable frequency. This work is dependent on the depth of penetration and the size of the part. The work is so heated that when the parts are brought to temperature they can be welded by the application of a suitable pressure. One of the members can be fixed and the other moved.

This method of welding has been particularly suitable for the welding of high-pressure boiler tubing, for the welding of railroad rails in the field, and for the butt welding of certain types of large cylindrical objects. Equipment has been furnished to weld transmission pipe in the field. A weld made in this manner is usually much faster than a weld made by a conventional electrical process.

Longitudinal Butt Welding. The longitudinal butt weld can be accomplished by induction on such mechanical equipment as a tube or pipe mill where hot-rolled skelp can be formed and welded continuously without the necessity of edge cleaning such as is required for conventional resistance welding. Of course, it is also possible to perform this type of welding by either induction resistance or high-frequency resistance welding.

21.14.4 Electron-Beam Welding

A method of welding, particularly in vacuum, has been developed which incorporates the utilization of an electron beam or beams for the welding of very thin and very narrow and unusual metals. This technique has been adapted to thin sections of conventional materials where it is desirable to obtain a weld by heating a very narrow section of the parent metals and obtaining a bond as strong as the parent metal. Commercial units for this type of welding are available, and the units are applicable to certain production operations as well as to laboratory welding.

21.14.5 Electroslag and Plasma Welding

Electroslag welding is a welding process which relies on a molten slag which melts the filler metal and the surface of the work to be welded. The weld pool is shielded by this molten slag which moves along the entire cross section of the joint as welding progresses. The electrically conductive slag is maintained in a molten condition by its resistance to the current which flows between the electrode and the work.

This type of welding appears particularly suitable to replace multipass welding techniques. The process is especially applicable for welding plates ranging in thickness from 1¼ to 18 in.

This method may require welding current of the order of 55 to 65 A for ⅛-in-diameter electrodes; however, the range can vary depending on the material to be welded from 400 to 900 A.

The welding voltage required is normally in the neighborhood of 40 to 55 V for ⅛-in-diameter electrodes.

The electroslag method has been widely used in Europe for some time. Its use should be considered when the welding application involves large structural members, pressure parts in nuclear reactors, and high-pressure boiler downcomer tubes. It could be used for forged crankshaft welding. It would be advisable in some instances to use it on press frames. It is suitable for cylinders for hydraulic presses, for turbines and alternator shafts, for rolling-mill frames, and for ship hulls.

Plasma is a field for electric heating now becoming commerically usable. There are basically two types of plasma heating: (1) arc plasma and (2) induction plasma.

Plasma heating involves the ionization of gases and develops extremely high temperatures.

Arc plasma heating has been used primarily for rapid melting of small quantities of material, both metals and refractories. It can be employed for cutting metals and is being used experimentally in investigations involving extremely high temperatures. A dc arc is generally used in obtaining the plasma.

Induction plasma heating is developed generally at frequencies of about 4 MHz. Experimentally, frequencies of 8 to 10 MHz have been utilized, and investigations are proceeding at many other frequencies. It is reported that plasma temperatures up to 20,000°F have been reached. Various gases are required, depending on the results desired. Gases commonly used are hydrogen, helium, argon, and nitrogen, and some work had involved the use of air.

21.14.6 Pressure Welding

Also called diffusion bonding or diffusion welding, the pressure-welding process is described as an atom-to-atom bond between two surfaces via intimate contact between their oxide-free surface areas. The surfaces must also be free of contaminants. To make the bond effective, the atoms on one surface must be brought, under pressure, within a few angstroms of the atoms of the second surface.

In some cases, interlayers of foil are necessary to achieve an adequate bond, acting as barrier layers between incompatible materials, to allow bonding of materials more readily welded at low temperatures, and to overcome problems with surfaces that have insufficient mobility to come into close contact.

Bonding is done in the solid state, with times and temperatures depending on the materials used. The method is advantageous because (1) it is applicable to a wide variety of joints; (2) many bonds can be made in one operation; and (3) bonds can be made in particularly hard-to-get-at places. At present, the disadvantage of diffusion welding is that joint assessment is difficult to make, but as technology improves, it is thought that, in the future, this type of welding may replace arc welding for major structural components such as nozzle and flange assemblies.

21.15 AIR CONDITIONING AND REFRIGERATION

21.15.1 Air Conditioning

The Air-Conditioning System. As described by the American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE), an air-conditioning system maintains desired environmental conditions within a space. Air-conditioning systems are categorized by how they control cooling in the conditioned area. Typical central air-conditioning system applications include (1) spaces with uniform loads, (2) small spaces requiring precision control, (3) multiple systems for large areas, (4) systems for complete environmental control, and (5) a primary source of conditioned air for other subsystems. Spaces with uniform loads are generally those with relatively large open areas and small external loads, such as theaters, auditoriums, department stores, and the public spaces of many buildings. Isolated rooms within a larger building with cleanrooms and computer facilities will typically require precision control. Large buildings such as factories, airplane hangers, shopping malls, office buildings, and hospitals require multiple systems.

Environmental Concerns. Owners of many medium-size and most large buildings find it more economical to cool their buildings by distributing chilled water to remote air-handling units than use large space-consuming air ducts from a central air handler. As a result, electric centrifugal chillers are used extensively for cooling commercial buildings. These chillers use chlorofluorocarbons (CFCs) as a refrigerant, and represent a class of extremely stable chemical compounds that are non-toxic and inexpensive. International agreements have been made to phase out use of these chemicals, primarily as a response to the concern that the CFCs may be contributing to the depletion of the ozone layer in the upper atmosphere and to global warming. Other refrigerants are available that limit or eliminate this problem, hydrochlorofluorocarbons (HCFCs), which have one or more hydrogen atoms in their structure and therefore decompose much more readily in the lower atmosphere, and hydrofluorocarbons (HFCs), which contain no chlorine and have little impact on stratospheric ozone.

Substitute refrigerants are available for many CFCs used in chillers. HCFC-123 is the present substitute for CFC-11 (R-11) in chillers, and for CFC-12 (R-12) chillers the substitute is HFC-134a. But neither refrigerant is a "drop in" substitute on existing equipment. Only in cases where equipment is relatively new and is specifically designated "compatible" with the new refrigerants can the substitution be easily made. In most cases, however, new equipment must be purchased that has been designed to use these substitutes. Availability of this equipment is increasing, and the new units can be supplied with performance equal or very close to their CFC counterparts.

Alternatives to electric chillers include steam and gas direct gas-fired absorption chillers, and gas engine-driven chillers. In a specific installation, the engineer must be careful to consider all costs with each alternative and compare options on an equitable, life-cycle cost basis.

Psychrometric Formula. A major step toward the development of air conditioning as a science was made by Willis H. Carrier with the derivation of the rational psychrometric formula (*Trans. ASME*, 1911) which established the relationship between dry-bulb and wet-bulb temperatures, both easily ascertainable, and other important properties of the air such as density, relative and absolute humidity, vapor pressure, and total heat, the exact valuation of which is essential to the solution of problems involving the exchange of heat and humidity between air-conditioning apparatus and conditioned spaces and their contents. The rational psychrometric formula is the basis for numerous psychrometric charts.

Psychrometric Terms

Dry-bulb temperature is the commonly observed temperature of the air as indicated by a conventional thermometer not affected by the relative humidity of the air.

Wet-bulb temperature is the temperature of adiabatic saturation of air or the temperature at which air would normally saturate without any change in its heat content. Wet-bulb temperature is, therefore, also a measure of the total heat content of air; that is, at a given wet-bulb temperature the total heat is constant regardless of changes in the dry-bulb temperature.

Relative humidity within the normal range is the percent ratio of the weight of water vapor actually contained in a unit of space to the weight of water vapor that the same unit of space would contain if fully saturated at the same dry-bulb temperature.

Absolute humidity is the actual weight of water vapor in air at any given condition and is usually expressed in grains per cubic foot or grains per pound of dry air.

Dew point is the temperature of saturation of air, or the point beyond which any further reduction in temperature would result in condensation.

Effective temperature is an arbitrary index of the degree of warmth or cold felt by the human body in response to temperature, humidity, and air motion. Experimentally determined, the scale has been fixed by the temperature of saturated air which reproduces the sensation of warmth felt at other combinations of conditions.

Sensible-heat effect is the exchange of heat between air and its surroundings or apparatus that changes the dry-bulb temperature without affecting the moisture content. It is convenient to consider sensible heat as dry heat.

Latent-heat effect is the exchange of moisture between air and its surroundings or apparatus by evaporation or condensation resulting in a change in the moisture content of the air with consequent change in the total latent heat of vaporization represented by the moisture in the air. It is convenient to consider latent heat as humidity.

Comfort Cooling. Investigations carried on under the auspices of the American Society of Heating and Ventilating Engineers (ASHVE), now ASHRAE, indicate that effective temperatures varying from 70 to 73°F are acceptable inside conditions for peak summer weather depending on length of occupancy and for normally clothed, normally active commercial workers. Table 21-8 gives approximate dry-bulb temperature and relative-humidity combinations within normal range which produce effective temperatures within the stated range with air motion 15 to 25 ft/min.

Inside design conditions of 78 to 80°F dry bulb with 45% to 55% relative humidity, depending on the character of occupancy, are accepted as standard inside design conditions against peak outside conditions for commercial comfort-cooling applications in almost all parts of the United States.

It is seldom desirable to design for inside temperatures in excess of 80°F. It is generally impracticable to reduce the humidity sufficiently to produce effective temperatures within the desired range at dry-bulb temperatures much in excess of 80°F.

TABLE 21-8 Inside Design Conditions

	Relative humidity, %		
	45	50	55
Effective temperature	Dry-bulb temperature, °F		
73	79½	79	78
72	78½	77½	77
71	77	76½	76
70	76	75	74½

Note: $t_{°C} = (t_{°F} - 32)/1.8$.

Industrial-Process Air Conditioning. In industrial or process air conditioning, inside design conditions are frequently dictated entirely by the product requirements. In confectionery plants where chocolate is handled, it is usually not desirable to have temperatures in excess of 65°F (dry bulb), which is far below the comfort range. In chemical or pharmaceutical plants where deliquescent chemicals are handled, lower-than-average humidities may be required for satisfactory production. Many products can be satisfactorily processed at widely varying conditions as long as the condition is maintained constant and the process adjusted to it.

Where conditions are established by product requirement, design is usually directed toward maintenance of satisfactory working temperatures. Investigations to determine the limits for optimum health and efficiency of the industrial workers indicate that conditions should not exceed 80°F effective temperature, that is, approximately 82.8°F, dry bulb at 80% RH, 84.4°F D at 70% RH, 86.2°F D at 60% RH, 88.2°F D at 50% RH, and 90.3°F D at 40% RH. Lower effective temperatures are required where perspiration must be prevented to avoid contamination or staining of products during handling.

There has been a marked increase in the use of complete air conditioning, including refrigeration, even in industrial operations where atmospheric control is not required by the process involved. An example is the trend toward larger manufacturing areas under single roofs, where windows cannot be depended on for light or ventilation. In many cases it has proved less expensive in both first cost and operating cost to provide full air conditioning with refrigeration rather than ventilation.

Winter Heating. Winter heating systems for normally occupied interiors are almost universally designed to maintain 70°F inside, with outside conditions assumed at 10°F above the lowest recorded temperature for the district in question.

Maximum Outside Design Conditions. These are usually taken at values which will not exceed 2.5% of the total hours during June, July, August, and September. Accepted maximum design dry-bulb (D) and wet-bulb (W) temperatures in degrees Fahrenheit for larger cities representative of various climatic divisions of the United States are given in Table 21-9.

Cooling Load. Calculation of the cooling load involves consideration of all possible internal as well as external heat gains not only for transmission and infiltration but also from sunlight radiation.

Solar Heat. Sunlight radiation for 40°N latitude on July 21 attains maximum values in Btu per square foot per hour transmitted by unshaded glass for variously oriented surfaces as follows: east, 200 at 8 A.M.; south, 91 at 12 noon; west, 200 at 4 P.M.; horizontal 258 at 12 noon (*ASHRAE Guide and Data Book*). Major consideration must be given to sun-exposed glass. Of the total impinging radiation, the amount passing through the net area of the window is approximately as follows: unshaded, 100%; half covered by buff shade, 70%; fully covered by light-finish venetian blinds, 70%; equipped with canvas awning, 30%. Although sunlight-radiation effect on walls and roofs can be calculated, allowance is usually made for this effect by calculating heat transmission at a temperature above the design outside temperature. A substantial temperature head (usually 50°F) should be allowed in calculating transmission through the roof surface due to long hours of exposure to the sun. Heat from sunlight contributes only to the sensible-heat load.

Heat Transmission. Transmission of heat through the structural surfaces of buildings or spaces is directly proportional to the temperature difference between the two sides of the surface and to the thermal conductivity of the material or materials forming it. Overall coefficients for a number of common structural elements are given in Table 21-10. The coefficient U is expressed in Btu per hour

TABLE 21-9 Maximum Summer Design Conditions

City	D	W	City	D	W	City	D	W
Boston	88	74	Cleveland	89	75	Seattle	80	65
New York	93	76	Chicago	91	76	San Francisco	77	62
Philadelphia	90	77	St. Louis	94	78	Los Angeles	90	70
Washington, D.C.	92	77	Atlanta	92	77			
Jacksonville	94	79	Birmingham	94	78			
New Orleans	91	80	Dallas	99	78	Denver	90	64
Houston	94	80	Phoenix	106	76	Salt Lake City	94	66

Source: *ASHRAE Guide and Data Book.*

TABLE 21-10 Coefficients of Heat Transmission *U*

Surface composition	
Single window	1.13
Double window	0.56
12-in brick wall, no interior finish	0.35
12-in brick wall finished with 3/4-in plaster on metal lath, furred	0.25
10-in hollow-tile wall, stucco exterior, no interior finish	0.33
10-in hollow-tile wall, stucco exterior finished with 3/4-in plaster on metal lath, furred	0.24
10-in concrete wall, no interior finish	0.61
10-in concrete wall finished with 3/4-in plaster on metal lath, furred	0.36
Wood siding, 1-in sheathing, studs, 3/4-in plaster on metal lath	0.26
Wood siding, 1-in sheathing, studs, 3/4-in plaster on metal lath, 3-in rockwool fill	0.07
Double partition, metal lath and 3/4-in plaster on both sides of studding	0.39
4-in hollow-tile partition, plastered on both sides	0.37
Yellow-pine flooring (25/32 in) on joists, no ceiling	0.45
Yellow-pine flooring (25/32 in), 3/4-in plaster on metal-lath ceiling below joists	0.31
6-in concrete slab floor, 1/4-in asphalt tile, no ceiling	0.60
6-in concrete slab floor, 1/4-in asphalt tile, 3/4-in plaster on metal lath hung or furred ceiling	0.38
1-in wood flat roof, built-up roofing, no ceiling	0.48
1-in wood flat roof, built-up roofing, 3/4-in plaster on metal-lath ceiling below joists	0.33
1-in wood flat roof, 1-in rigid insulation under built-up roofing, no ceiling	0.21
1-in wood flat roof, 1-in rigid insulation under built-up roofing, 3/4-in plaster on metal-lath ceiling below joists	0.17
4-in concrete-slab flat roof, built-up roofing, no ceiling	0.70
4-in concrete-slab flat roof, 1-in rigid insulation under built-up roofing, no ceiling	0.24
4-in concrete-slab flat roof, built-up roofing, 3/4-in plaster on metal-lath ceiling, hung or furred	0.41
4-in concrete-slab flat roof, 1-in rigid insulation below built-up roofing, 3/4-in plaster on metal-lath ceiling, hung or furred	0.18

Source: *ASHRAE Guide and Data Book.*
Note: 1 in = 25.4 mm.

per square foot per degree Fahrenheit temperature difference across the surface and is based on an outside-surface coefficient for 15-mi/h wind velocity.

Total heat transmission through any exposed building surface may be computed by the formula

$$H = A \times U(t_o - t_i) \tag{21-21}$$

where H = sensible heat in Btu per hour, A = area of the surface element in square feet, U = coefficient of heat transmission of the surface, and t_o and t_i are, respectively, the outside and inside temperatures in degrees Fahrenheit. This formula is used for computing heat loss from a space as well as heat gain to it.

Heat from Occupants. Every occupant of a conditioned space contributes both sensible and latent heat to the atmosphere; the relative proportions is dependent on the dry-bulb temperature and the total quantity on the state of activity of the individual (see Table 21-11).

Only the sensible heat(SH) fraction affects the temperature of the conditioned space, since latent heat (LH) from evaporation of body moisture appears as vapor at room temperature.

Electrical Heat. The heat equivalent of the electrical input to all lights and electrical heating or power appliances used within a conditioned space contributes to the atmospheric heat which must be removed by the air-conditioning apparatus. In general practice, the approximate heat equivalent of

TABLE 21-11 Heat Loss from Average Person, in Btu per Hour, as Sensible Heat (SH) and Latent Heat (LH)

		Dry-bulb temperature, °F					
		70		75		80	
		Segregated heat loss, Btu/h					
Condition of activity	Total heat loss, Btu/h	SH	LH	SH	LH	SH	LH
Seated at rest	400	300	100	260	140	220	180
Light work	660	350	310	290	370	220	440
Moderate work	850	430	420	360	490	270	580

Note: 1 Btu/h = 0.293 W; $t_{°C} = (t_{°F} - 32)/1.8$.

3400 Btu/h · kW input to lights and appliances and 2500 Btu/h · hp of motor capacity is used in calculating heat gain from this source. In considering heat from appliances and motors, load factor should be carefully investigated, since appliances or motors seldom run continuously at full load. Unless appliances are used for water heating or evaporation, the entire load from electrical sources will be sensible heat.

Process Heat. Heat from gas appliances is the most common form of process heat. Average heat input to common gas appliances in Btu per hour is as follows: 14-in coffee urn, 7500; 12-in coffee urn, 5000 (*ASHRAE Guide and Data Book*). Heat output will be approximately 75% sensible heat and 75% latent heat. Consideration should be given to load factor of appliances.

Ventilation and Infiltration. Air quantity for ventilation purposes is usually based on the number of occupants of the conditioned space. Ventilation is required chiefly to overcome normal odors. Average accepted practice calls for the introduction of at least 10 ft^3/min of outside air for each nonsmoking occupant and 20 ft^3/min for each smoking occupant. For commercial interiors, such as restaurants and drugstores, where moderate smoking occurs, 15 ft^3/min per person is usually allowed. In theaters, where ceiling heights are usually liberal, where there is no smoking, and peak occupancy seldom occurs for long periods, an allowance of 7½ ft^3/min per peak occupant is considered ample. Air positively exhausted from a space must be deducted from the positive ventilation supply in considering possible infiltration. If the exhaust exceeds the supply, infiltration will naturally be induced. Air for ventilation should be first brought to the conditioning apparatus and conditioned before introduction to the conditioned space.

Infiltration varies with the construction, exposure, and size of a conditioned area as well as with door traffic and location of doors. The following quantities are suggested as a general guide for buildings of average construction: cubical contents 0 to 5000 ft^3, one complete change of volume in 30 min; 5000 to 50,000 ft^3, 40 min; 50,000 to 100,000 ft^3, 60 min; 100,000 to 200,000 ft^3, 90 min; over 200,000 ft^3, 120 min. Infiltration in cubic feet per minute for load calculation and comparison with positive supply in cubic feet per minute may be determined by dividing cubical contents of the space by the time factor for that space magnitude.

The sensible-heat component of ventilation or infiltration may be computed approximately from the formula

$$H_s = Q \times 1.07(t_o - t_i) \qquad (21\text{-}22)$$

where H_S = sensible heat in Btu per hour, Q = cubic feet per minute of air, and t_o and t_i = outside and inside temperatures in degrees Fahrenheit, respectively.

The latent-heat component of ventilation or infiltration may be computed approximately from the formula

$$H_L = Q \times 0.675(h_o - h_i) \qquad (21\text{-}23)$$

where H_L = latent heat in Btu per hour, Q = cubic feet per minute of air, and h_o and h_i = outside and inside absolute humidities, respectively, expressed in grains per pound of dry air (see a psychrometric chart).

Only infiltration air entering the conditioned space directly contributes to the internal heat load. Ventilation air-conditioned before introduction to the space contributes to the load upon the conditioner and cooling apparatus only.

The Internal Sensible-Heat Load. This will be the summation of the sensible heat from transmission, sunlight radiation, occupants, lights, electrical devices, process heat, and infiltration. The *internal latent-heat load* will be the summation of the latent heat from occupants, process heat, and infiltration. A factor of safety of 10% is often added to these figures to allow for inaccuracies and, in the case of sensible-heat load, to cover the power input for air circulation. The *total heat load* on the conditioner or cooling apparatus will be the sum of the internal sensible- and latent-heat loads plus the sensible- and latent-heat load of ventilation air. Grand-total heat is usually expressed in tons of refrigeration effect.

Air Quantity for Sensible-Heat Absorption. The sensible-heat load usually serves as the first guide to the selection of proper air quantity. The specific heat of air within the normal air-conditioning range is approximately 0.241 Btu/lb · °F. One Btu will raise 4.16 lb of air 1°F, or approximately 56 ft^3 of air 1°F. The air quantity required to absorb a given sensible-heat quantity may be determined by the formula

$$Q = \frac{H_S}{60} \times \frac{56}{t_i - t_d} = \frac{H_S}{1.07(t_i - t_d)} \tag{21-24}$$

where Q = air quantity in cubic feet per minute, H_S = sensible heat in Btu per hour, t_i = space temperature in degrees Fahrenheit, t_d = delivery temperature of air into the conditioned space in degrees Fahrenheit, 60 = minutes per hour, and 56 = cubic feet per Btu per degree Fahrenheit temperature rise. Modern diffusion-type outlets will permit satisfactory delivery of air into spaces of 10- to 12-ft ceiling at 60°F. With higher ceiling heights and good conditions, air may be introduced at 55 to 50°F.

Latent-Heat Absorption (Dehumidification). For the quantity of air selected to meet the sensible-heat load by the method just suggested, the moisture content at which the air must be introduced into the space to absorb the internal latent-heat load may be approximately calculated by the following formulas:

$$h_r = \frac{h_L}{Q \times 0.675} \qquad \text{and} \qquad h_d = h_i - h_r \tag{21-25}$$

where h_r = humidity deficiency in grains per pound of dry air; h_d and h_i = absolute humidity in grains per pound of dry air at delivery condition and within the room, respectively; H_L = internal latent heat in Btu per hour; Q = air delivery in cubic feet per minute; and the factor 0.675 is calculated from the average density of air, the average latent heat of moisture in Btu per grain, and rate correction. The delivery-moisture condition thus calculated, together with the delivery temperature, will permit graphical determination of the required delivery-air conditions from a psychrometric chart. *Dehumidification* is usually accomplished by condensation of moisture in the process of cooling the air.

Cooling-Dehumidifying Apparatus. Two types of apparatus are used for cooling and dehumidifying: (1) the spray-type dehumidifier and (2) the surface cooling coil. In either, cooling and dehumidification may be accomplished by cold fluid at the proper temperature from any source, for example, cold water from wells or ice tanks or water cooled directly by refrigeration apparatus or by the direct expansion of refrigerants in cooling coils.

Spray Type. In the spray-type dehumidifier, air to be conditioned is passed through a spray chamber into which a large volume of finely divided water is introduced. Spray-type dehumidifiers are of two general types: (1) those in which water cooled externally is introduced at a reduced temperature and (2) those in which water is sprayed over coils cooled by fluid from an external source.

Spray-type dehumidifiers have the advantage of permitting evaporative cooling of cleaning the air of siliceous dust, and of absorbing some gases and odors.

Surface Cooling Coils. Surface cooling coils are widely used for cooling and dehumidification. Cooling of surface coils is usually accomplished by direct expansion of refrigerant though frequently cold water or brine is used.

Common designs include (1) pierced-plate types in which the tubes pass through closely spaced thin metallic plates, (2) helical extended surface in which a ribbon of metal crimped or slit to accommodate curvature is wound in a helix about the central tube, and (3) the type in which fins are extruded directly from the tube body by rolling.

Heating Surface. Heating surface is generally similar to cooling surface in construction. Extended cast-iron surface is still used to some extent. Steam is an accepted heating medium, although the use of hot water is increasing, as it facilitates control.

Humidification and Humidifiers. In many comfort cooling applications and some industrial-process operations which have low winter heat loads and require only moderate humidification during the winter, this is accomplished by small sprays which introduce finely divided water droplets into the air steam usually directly before or after the heating coil so that evaporation is stimulated. Pan-type humidifiers in which water is evaporated into the air steam by steam coils or direct steam sprays may also be employed. It should be noted, however, that these procedures add to the loading within the space and therefore should be employed only where there is a deficiency in heat gain or a heating requirement.

In large commercial or industrial operations employing spray-type air washers for summer cooling and dehumidification, humidity control is usually achieved by maintaining a constant dew-point temperature and then allowing the air to rise to the dry-bulb temperature, resulting in the desired final relative humidity in absorption of heat gains within the space. It should be noted that the majority of industrial and many large commercial applications involve internal loading from power, lights, etc., which require the application of cooling during all seasons of the year when there is normal occupancy or operation. During winter this cooling may be obtained by outside air, as described subsequently for evaporative cooling, but there is no heating requirement during normal service.

The major shortcoming of the dew-point control method for humidity maintenance lies in the fact that there must be a fixed relationship between the maintained dew point and the final dry-bulb condition within the space to achieve the desired relative humidity. Thus there is a fixed relationship between the load and the temperature rise in supply air which may be taken to absorb this load, and this relationship may result in the requirement for excessively high circulating-air quantities, particularly if high relative humidities must be maintained as is common in many textile-mill operations.

Accordingly, in such operations where high relative humidities must be maintained with high continuous internal heat loads, common practice is to employ a *split system* in which the dew point of the circulating air is maintained at an arbitrarily low level which would normally result in insufficient humidification after absorption of the heat load, but the humidity is increased within the space by the direct application of water droplets from compressed-air or motor-operated mechanical atomizers. The atomized moisture introduced directly into the space is in water form and in the process of evaporation not only increases the relative humidity but also absorbs sensible heat from the space in the process of being converted from water to vapor at the room condition. The moisture thus introduced does not add anything to the heat load within the space and further, by the process noted above, increases the sensible-heat-absorbing capacity of the primary supplied air, thus making it possible to perform the required cooling with a substantially lower volume of circulating air than would be needed for the straight dew-point method.

Atomizers are also frequently used as prime means for humidification in areas such as textile mills. Employed in this fashion or with the central-station system, they also afford an excellent facility for direct humidity control in zones within a major area.

Evaporative Cooling. During intermediate seasons in normal climates or in climates such as prevail in dry mountainous country where temperatures are frequently excessive but humidities continuously low, evaporative cooling systems are very effective. Air is saturated adiabatically and thus assumes the wet-bulb temperature. Sensible cooling is thus done at the expense of increased humidity. Calculations for evaporative cooling systems are identical with those for other cooling loads except that 100% outdoor air is used.

Commercial evaporative coolers are similar to spray-type washers or humidifiers or spray-type dehumidifiers. The spray water is continuously recirculated by means of a pump.

Fans for Air Conditioning. Many different types of fan are used in HVAC systems, each of which has certain advantages. Centrifugal fans generate pressure by accelerating air outward from a fan wheel and then converting this kinetic energy to pressure. Centrifugal fans are able to generate relatively high pressures, but tend to be large and heavy with high WR^2 loads. Axial fans, on the other hand, generate pressure using the lift created by the fan blades much like a propeller. Axial fans generate high flowrates at low pressures and have low rotating-mass characteristics. In general, axial fans have to rotate much faster than do centrifugal fans to generate comparable pressures and flow rates. As a result, they tend to generate higher airborne noise levels. However, their lower weight and space requirements make axial fans preferable to centrifugal fans for many ventilation applications. Their high rotating speed requirements are also suitable for direct drive by 4-pole and 6-pole induction motors. By eliminating the belt, these fans are more compact and energy-efficient because of the avoided losses in the belt drive.

In residential applications, the forward-curved-blade centrifugal fan is commonly used because of its ability to generate pressure and flow at relatively low operating speeds. This low-speed operation allows forward-curved fans to generate low noise and structural vibrations. However, in general, forward-curved fans are relatively inefficient, and are not used in larger applications where long operating times can incur high energy costs. Residential HVAC fans are usually powered by single-phase, multispeed motors.

In many commercial applications, backward-inclined centrifugal fans are used to move large amounts of air. When equipped with airfoil-shape blades, these fans can achieve high operating efficiencies. In systems that are distributed over a large floorspace, smaller fans are installed throughout the system ductwork to boost air pressure to the points of delivery.

In many industrial applications, fans also serve to exhaust contaminants and to provide convective cooling by circulating air within a workspace. In simple applications, basic propeller fans are installed without ductwork in rooftops to remove inside air. In more complex applications, fan selection is based on the size, configuration, and pressure requirements of the ventilation system, air properties, contaminant levels, and number of operating hours. Radial-blade and radial-tip centrifugal fans are used in ventilation applications with high particulate levels. The shape of these blades prevents the accumulation of contaminants on the blade surfaces.

In emergency ventilation systems, fan operation is relatively infrequent; consequently, tubular axial fans are often used because of their low starting torque requirements and their space saving features. Most industrial HVAC fans are driven by NEMA Design B polyphase induction motors. Service factors for centrifugal and axial fans are typically 1.10 to 1.15.

Fans are commonly connected to V-belt drives. By using different pulley sizes between the motor and the fan, belt drives allow fans to rotate at slower speeds than their motors. Most fan motors are 2-pole, 4-pole, or—in some cases—6-pole, which rotate at 3600, 1800, and 1200 rpm, respectively. In many HVAC applications, multispeed motors allow the selection of two of these speeds. By using a larger pulley on the fan than the motor, the fan rotates at a lower speed, reducing noise and vibration levels. Speed reduction ratios range between 1–1.5 and 1–4.

Adjustable-Speed-Drive (ASD) Applications. In ventilation systems, the larger opportunities for energy improvement are through the use of variable-speed drive fans and variable-air-volume systems. *Variable-air-volume* (VAV) systems rely on the use of dampers at the point of air delivery to control the amount of heating or cooling air entering a space. VAV systems are improvements over previous system designs such as dual-duct and reheat systems. Dual-duct systems mix cool and warm air streams to achieve a desired delivery temperature. Reheat systems cool air to temperatures between 45 and 60°F, to remove moisture, then reheat it to the desired delivery temperature. Unlike VAV systems, which avoid the unnecessary heating and cooling of air, dual-duct and reheat systems are relatively energy inefficient. Adjustable-speed drives are well suited to operate with the circulating fans in VAV systems, can control the operating speed of a fan to regulate airflow, and can be equipped with feedback instrumentation that matches fan speed to system airflow requirements. By slowing fan speed during periods of low system demand, ASDs minimize the amount of excess flow energy imparted to the system.

Variable-frequency drives (VFDs) are the most common type of ASD. VFDs rely on rectifiers and inverters to adjust the frequency of the power delivered to the motor. VFDs can be combined with direct-drive units to offer attractive advantages in speed control, efficiency, and space savings. VFDs are widely used in commercial and industrial applications, and are included with up to 10% of all new fan installations.

Automatic Control. Most control devices depend on physical, thermal, or hygrostatic force for their operation. For example, thermostats frequently are operated by the distortion of bimetallic strips caused by the unequal expansion of the metals with increase in temperature or by the increase in pressure with rise in temperature of a liquid or gas contained in a closed element. Hygrostats generally depend on the expansion or contraction of human-hair strands with changes in humidity to perform the basic control operation, although silk, paper, or wood is used in certain applications. The thermostatic and hygrostatic elements usually operate other devices to perform the control operation.

Air-Conditioning Control Equipment. This equipment may be divided into three separate classes: self-contained, pneumatic, and electrical.

Self-contained devices are usually composed of thermostatic-fluid elements directly operating steam, water, or brine valves. These devices are not used where close control is necessary or where instantaneous tight shutoff is required. However, they are simple, cheap, and reliable and inherently give modulated control.

Pneumatic-control systems are those in which the thermostatic or hygrostatic element controls the rate of air leakage from a compressed-air system. The air pressure in the control system is used to operate diaphragm motors controlling valves and dampers. The spring loading of these motors can be varied so that within the given control-pressure range a whole sequence of separate valve and damper operations can be accomplished. Various types of relay and compensating arrangements are used to overcome inherent tendencies of such systems to override or "hunt" if set for high sensitivity or to give too wide a control range if set for low sensitivity.

The most common *electric thermostatic or hygrostatic controls* are of the *snap-action type*; that is, the condition change simply serves to cause the device to make or break an electric circuit, which in turn controls the major element or elements to be operated. Because of contact difficulties, line-voltage controllers generally use mercury-tube contactors, which limit the accuracy of control. All wiring to the control must be run in conduit. Low-voltage controls (24 V ac) which operate major motor drives or other elements through relays are more common; the light, open contacts permit accurate control, and open wiring (usually metal-clad) is allowed.

Electric controls employing thermocouples or resistance thermometers for sensing temperature conditions and various electrical leakage devices to determine humidity conditions are finding extensive application for air-conditioning control. However, for large-system work in the actuation of valves and dampers it is hard to find a satisfactory substitute for the pneumatic or hydraulic operator, so that electric or electronic control circuits usually act through other relays to actuate the major system elements by means of such operators. Nevertheless, there is an increase in the use of electric and electronic equipment because of the trend to centralizing main control observation, recording, and actuating points as a result of the complexity and extent of the systems being installed in large commercial and industrial operations.

Air-Conditioning Systems. *Central systems* are usually assemblies of air-conditioning apparatus designed to afford the necessary atmosphere control and may be arranged with (1) a single fan and single apparatus to care for one space; (2) two or more fans and a single apparatus to care for two or more zones; (3) a single fan blowing through warm and cool air chambers with mixing dampers delivering properly proportioned air to two or more zones; (4) a single apparatus and central fan delivering highly conditioned air to individual recirculating fans in two or more zones; or (5) a single central apparatus and fan with supplementary cooling or heating equipment in the ducts leading to two or more zones.

The wide application of air conditioning to multiroom, multistory commercial structures presents a number of problems which have resulted in the development of a number of new systems and techniques. Among systems in common use are the following:

1. The fan-coil system which employs individual units served with hot or cold water from a central circulating system or zone.
2. The high-pressure induction unit system where highly conditioned air at a high velocity is used to aspirate the flow of room air through localized heat-transfer coils supplied with hot or cold water from an external source, the primary air caring for ventilation and dehumidification while the secondary coil cares for the major sensible-heat gain or loss within the zone.
3. The low-pressure hot and cold plenum system in which each zone is supplied from a central unit or apparatus where warm and cold air are mixed as required to meet the zone demand.
4. The high-pressure double-duct system in which warm and cold air from central apparatus served through parallel trunk supply systems are mixed at a local distributing point in each zone through a sound-deadening mixing-valve device subject to control from the immediate zone.
5. The single-effect high-pressure interior system for substantial areas with relatively constant load characteristics with control for various sections by manual or automatic volume regulation, etc.

The Unit Air Conditioner. This is a factory-assembled system within a suitable casing of component apparatus necessary for atmosphere control which may include filters, sprays, cooling or heating coils, fan motor, dampers, etc. Unit air conditioners which include in the same casing the refrigerating system are defined as *self-contained unit air conditioners. Direct units* are those designed for location within the conditioned space. *Indirect or remote units* are those which are designed for location outside the conditioned space. *Free-delivery units* are those with fixed fan characteristics designed for use without air ducts.

Self-contained room air conditioners designed for mounting in a window or special opening through an outside wall have been highly developed as a plug-in appliance requiring only electrical connection. These units consist of filter and cooling coil, conditioned-air fan, and outlet and inlet grilles, together with refrigeration compressor, motor and drive, condenser air fan, and air-cooled condenser, the latter elements located outside the building wall.

These units are available with total heat-removal capacities in Btu per hour ranging from approximately 5000 to 24,000 and with nominal power input approximately 1 hp/10,000 Btu capacity. Units up through 8000 Btu capacity are generally suitable for operation on 115-V single-phase service. Larger units usually require 230-V single-phase service, but this is available for such use in most areas. Units of this type almost universally employ air-cooled condensers.

Self-contained commercial units with the refrigeration system built in as part of the factory assembly are manufactured in various sizes with nominal total heat-removal capacities ranging from 2 tons to in excess of 30 tons. Units in capacities of 2 to 10 tons are frequently employed as direct units within the space, and many employ water-cooled condensers. However, the current trend is to air-cooled condensers located external to the space. A number of designs are available in capacities ranging from 5 to 30 tons designed for roof mounting and complete with air-cooled condensers and in many instances with electric-resistance heating coils or automatic gas or oil-fuel burner systems for heating. Such units are sometimes used in multiple to take care of fairly substantial commercial or industrial areas in single-story construction.

Self-contained and factory-assembled condensing units complete with either water- or air-cooled condensers and suitable for connection to remote air-conditioning units are currently available in capacities of 2 to 75 tons.

Drives for Air-Conditioning Systems. There is no satisfactory substitute for electric-motor drive for the power auxiliaries of air-conditioning systems. Compression-refrigeration systems which are used in conjunction with the large majority of air-conditioning installations are primarily designed for electric-motor drive.

Fuel and Energy Conservation. Because of the rising cost of fuels and energy, and the rapid rate at which fossil fuels are being used, methods of energy conservation are being emphasized. Some of these include

1. *Use of computer programs* to determine the efficient energy use of the building or plant. Massive data relating to constants and variables, such as building occupancy, fuel consumption, and electrical characteristics of the building, are fed into the computer and compared with programs to conserve fuel. An analysis of the comparison shows which proposals are feasible and which are not.
2. *Automated monitoring* of key systems; for example, logic controllers can be set up on boilers to ensure efficient fuel utilization automatically.
3. *Heat-recovery systems*. Heat-recovery cooling units are growing in popularity; typical systems can be water-to-air or air-to-air types. Heat is recovered from areas of a building or plant that requires cooling and transferred to a portion of the plant that requires heating. Some systems recover flash steam for building heating, and still other systems can turn a dust-collection unit into a heating system by filtering the dust and particles through a hopper into drums for disposal, while redirecting the clean filtered air into the plant or building for heating.
4. *Solar air conditioning, heating, and cooling*. Solar collector arrays and reflector panels are located on a building's roof and are capable of regulating and storing the energy received from the sun for equal distribution during the day.

21.15.2 Refrigeration

Refrigeration Cycles. All practical refrigeration cycles produce heat removal by free evaporation in an enclosed chamber (evaporator) of a liquid (refrigerant) under pressure conditions that produce the desired evaporation temperature. The refrigerant liquid absorbs its latent heat of vaporization from the medium being cooled and in this process is converted to vapor at the same pressure and temperature. This vapor is conveyed to another chamber (condenser), in which the pressure is maintained at a level sufficiently high to permit condensation of the refrigerant by water or air at normal temperatures. The heat quantity abstracted in the condenser is the latent heat of condensation (or vaporization reversible) of the refrigerant fluid, together with the heat that has been added to the refrigerant in the process of conveying it from the evaporator pressure level to the condenser pressure level. The condensed refrigerant (liquid) is allowed to flow from the condenser, through suitable throttling valves, back to the evaporator to repeat the cycle.

Unit of Refrigeration. The standard unit of refrigeration is the ton, which is considered as heat removal at the rate of 200 Btu/min or 12,000 Btu/h or 228,000 Btu/24 h. This rate of cooling is about equivalent to the average cooling effect obtained by melting 1 ton of ice in 24 h at 43°F (latent heat of fusion of ice is 144 Btu/lb).

Refrigeration Systems. Practical refrigeration systems differ only in the method used to convey the refrigerant vapor from the evaporator, or low side, to the condenser, or high side. In closed refrigeration systems, three methods are used to accomplish this transfer: (1) mechanical compression or pumping of the refrigeration vapor, (2) chemical absorption of the refrigerant vapor at low pressure with subsequent transfer to the high side as solute in a solution, and (3) physical absorption of the refrigerant vapor at low pressure with subsequent transfer to the high side as adsorbed vapor in a solid. Related to systems (2) and (3) are absorption and adsorption dehumidification systems used in air-conditioning processes requiring drying of air without cooling it.

Compression-Refrigeration Cycle. In the compression-refrigeration cycle a mechanical compressor or pump is used to convey the refrigerant vapor from the (low) evaporator pressure to the (high) condenser pressure. Vapor is drawn from the evaporator, where it has been evaporated in performing the cooling work, and is compressed and delivered to the condenser where it is liquefied, its latent heat of vaporization being absorbed by the condenser-cooling medium, which is usually water or air at normal temperature. The liquefied refrigerant is collected in the bottom of the condenser or in a separate container called a "receiver" and from there is fed back to the evaporator through suitable throttling valves.

There are three types of mechanical compressor in common use: (1) the reciprocating compressor, (2) the rotary compressor, and (3) the centrifugal compressor. To these must be added the jet compressor, which, while not usually considered as a mechanical compressor, is quite similar in operation. The *reciprocating and rotary compressors* are positive-displacement compressors; that is, each cycle of the compressing member definitely opens a fixed volume into which fluid flows and then closes or occupies that same volume, forcibly expelling the fluid. The centrifugal or jet compressors depend for compression on the kinetic or velocity energy imparted to the fluid.

The *reciprocating compressor* is suited to the compression of relatively small volumes of fluid through a high-pressure range. The *rotary compressor* is suited to the movement of moderate fluid volumes through moderate pressure ranges. The *centrifugal compressor* is suited to the movement of large volumes of fluid through small pressure ranges. Ammonia, carbon dioxide, Freon-12, methyl chloride, and sulfur dioxide are almost always used with reciprocating compressors. Freon-21 is used with rotary compressors, while Freon-11, methylene chloride, and water are used with centrifugal compressors.

Refrigerants. The refrigerants which are used commonly enough to merit classification by the National Board of Fire Underwriters are listed in Table 21-12. The Freon group is now among the most widely used of all refrigerants. These halogenated hydrocarbons are chiefly derived from methane (CH_4) by the replacement of hydrogen molecules with chlorine and fluorine molecules and include F-11, F-12, and F-22.

Those listed as inflammable or in toxicity classification groups 1 to 4 (mildly to seriously toxic depending on concentration, exposure, etc.) include ammonia, butane, dichloroethylene, ethane, ethyl bromide, ethyl chloride, methyl bromide, methyl chloride, methyl formate, propane, and sulfur dioxide.

Those listed as nonflammable (or practically so) and less toxic than group 4 (nontoxic in normally possible concentrations) include carbon dioxide, dichlorodifluoromethane (F-12), monochlorodifluoromethane (F-22), dichlorotetrafluoromethane (F-114), monofluorotrichloromethane

TABLE 21-12 Comparison of Refrigerants
Refrigerating Data Book

Refrigerants	Displacement cfm per ton	Work, hp per ton	Coefficient-of performance	Pressure of saturated vapor, lb/in² (abs),* at saturation temperature 5°F	40°F	86°F	100°F
1. Anhydrous ammonia (NH_3)	3.44	0.989	4.76	34.27	73.32	169.2	211.9
2. Carbon dioxide (CO_2)	0.96	1.840	2.56	331.9	567.8	1043	
3. Dichlorodifluoromethane (CCl_2F_2)(F-12)	5.81	1.002	4.70	26.51	51.68	107.9	131.6
4. Monochlorodifluoromethane ($CHClF_2$) (F-22)	3.60	1.011	4.66	43.02	83.72	174.5	212.6
5. Monofluorotrichloromethane (CCl_3F)(F-11)	36.32	0.927	5.09	2.93	7.02	18.30	23.60
6. Methyl chloride (CH_3Cl)	5.95	0.962	4.90	20.80	42.60	95.50	119.0
7. Methylene chloride (CH_2Cl)	74.30	0.963	4.90	1.17	3.38	10.60	13.25
8. Sulfur dioxide (SO_2)	9.09	0.968	4.87	11.81	27.10	66.50	84.50
9. Water (H_2O)	476.70	1.125	4.10	. . . .	0.25†	. . . .	1.93†

*To obtain gage pressures, deduct 14.7 lb/in² atmospheric pressure. [1 lb/in² = 6.895 kPa; $t_{°C} = (t_{°F} - 32)/1.8$; 1 ton = 907.2 kg.].
†These pressures given in inches of mercury absolute.
Note: Figures are based on standard conditions and dry compression and represent the theoretical performance of the refrigerants for these conditions based on their thermodynamic properties. Figures for carbon dioxide and water are unfavorable under standard conditions, since 86°F condensing temperature is close to the critical temperature of CO_2 and 5°F suction temperature is below the freezing point of water.

(F-11), and methylene chloride (dichloromethane). Water is also used as a refrigerant and is, of course, considered as nontoxic and nonflammable.

Refrigerant Properties. Physical and thermodynamic properties of the nine most commonly used refrigerants and comparison of the performance of these refrigerants with the ideal (Carnot) cycle for standard conditions are listed in Table 21-12.

Ammonia is cheap and easily available; its efficiency is so high that it is the practical standard by which other refrigerants are rated. For these reasons it is widely used for ice manufacture and food processing, for example, ice-cream manufacture, meat packing, cold storage, and brewery refrigeration.

Freon refrigerants, which have satisfactory characteristics with respect to toxicity and flammability as well as stable operation characteristics within the normal range, have largely replaced carbon dioxide for air conditioning, etc., where toxic refrigerants present a public hazard.

Methyl chloride, sulfur dioxide, and Freon refrigerants find broad application in commercial and domestic refrigeration service.

Effect of Operating Conditions on Compressor Capacity and Power Requirement. The capacity of any refrigerating compressor depends on the mass of refrigerant vapor that it can convey from the low- to the high-pressure condition. The power requirement of the compressor depends on the mass of vapor conveyed and the pressure differential between the evaporator and the condenser. For this reason the capacity and power requirement of any given compressor operating at constant displacement vary widely depending on the temperature and pressure conditions of the suction and discharge vapor.

It should be noted that although the unit power requirement is much greater for low suction temperatures, the capacity of a constant-displacement machine is greatly reduced owing to reduced density of the vapor and reduced volumetric efficiency, so that the gross power requirement may be far below that at higher suction temperatures. For this reason it is general practice to operate belt-driven compressors at higher speeds when used at lower suction pressures to obtain greater displacement and capacity within the safe power rating of the machine. Care must be exercised in selecting equipment to be sure that safe power limitations will not be exceeded during start-up periods when abnormally high temperatures may be encountered owing to high refrigerant temperatures. This is a very important consideration in the design of ultra-low-temperature systems, especially those employing multiple compressors in cascade arrangement. Within the normal range of condensing temperatures, the change in vapor density is not so pronounced, and the chief effect of departure above the design condensing temperature is a reduction in the output of the compressor.

Evaporators. The largest classification of evaporators for direct air cooling is surface-cooling coils. For cooling water or brine, simple arrangements of submerged pipe coils or plate surface are sometimes used for small applications, especially where refrigeration storage is desirable. More widely used with reciprocating-compression equipment is the so-called dry-expansion liquid cooler in which refrigerant is expanded into the tube surface within a shell and water passed under forced circulation across the tube surface as directed by baffles. Some larger reciprocating-compressor installations and all large-capacity centrifugal refrigeration systems employ a conventional shell-and-tube surface, with forced circulation of water through the tubes, which are mounted within a pool or evaporating refrigerant within the shell surface.

The majority of such applications employ integrally finned tubing.

Expansion Methods. Two general types of refrigerant expansion are employed: (*a*) dry expansion and (*b*) flooded expansion.

In the *dry-expansion system* the refrigerant is introduced into the evaporator through an expansion valve or pressure-reducing valve and makes one pass through the evaporating surface going to the compressor or absorber suction line. Since any liquid refrigerant which leaves the evaporator represents a loss of cooling effect, care is taken to assure only dry gas leaving the evaporator. The constant-pressure valve has been largely supplanted by the thermal-expansion valve, which through the use of the self-contained thermostatic-valve principle controls the flow of refrigerant to give a constant superheat at the outlet of the evaporator regardless of the load. Dry-expansion systems are widely used with expensive refrigerants, since only a minimum refrigerant quantity is required to feed the evaporator properly. The dry-expansion system also permits feeding evaporators from the top so that any oil carried in the refrigerant will flow through the evaporator and be conveyed back to the compressor.

The flooded-expansion system maintains through float-valve control a constant level of refrigerant in the evaporator. In coil-type evaporators this is accomplished by means of a surge chamber external to the coil, usually located above the top and bottom of the coil. Refrigerant is fed into the surge chamber, which is connected with the top and bottom of the coil and to the suction line. As liquid refrigerant passes up through the evaporator coil, its density is reduced owing to the formation of vapor bubbles, so that the refrigerant circulates rapidly through the coil to the surge chamber, where the gas bubbles are released to the suction line. The major disadvantages of the flooded system are the relatively large refrigerant charge required and the necessity of providing means for oil removal when refrigerants that mix readily with oil are used.

Condensers. The function of the condenser is exactly the reverse of the evaporator, that is, to afford a rapid transfer of heat from the condensing refrigerant to the cooling medium. For this reason there is a great similarity between condensing and evaporating equipment. Air-cooled condensers are frequently used with small compression systems which are located in well-ventilated places and where the difficulty of obtaining water supply and drain connections is not commensurate with the operating economies realized through the use of water cooling. Air-cooled condensers are similar in construction to surface-cooling coils.

Because of the effect that high condensing pressures have upon the economy of operation of refrigeration plants, water-cooled condensers are widely used. It is customary to assume the condensing requirement of the average compression plant to be 15,000 Btu/(ton)(h), of which 12,000 Btu represents the cooling work, and 3000 Btu the energy applied in heat pumping. This amount of heat can be absorbed by 1 gal/min of water rising through 30°F, 2 gal/min through 15°F, or 3 gal/min through 10°F. The design of the condenser and the amount of water required will depend on the available water temperature and the desired condensing temperature and pressure.

Water Conservation. The increased use of tap water for condensing purposes in air-conditioning and refrigeration installations has seriously taxed the *water-supply* and *water-disposal* systems of many cities, with the result that restrictions in the form of increased water rates for condensing purposes, taxes on water disposal, or outright prohibition of the use of water for condensing purposes have been effected by many municipalities. This has resulted in the development of equipment for the conservation of water. The familiar *cooling tower* has been redesigned to meet requirements of downtown buildings as to size and appearance, and the spray pond and atmospheric cooling tower find broader application in the industrial field. With spray ponds and atmospheric towers usually 5 gal/(min) · (ton capacity) is allowed with water temperature entering the condenser at 10°F above the maximum outside wet-bulb temperature. With forced-draft towers usually 3 gal/min · ton will suffice with the same temperature limits.

Thoroughly accepted in this field is the *evaporative condenser*, which consists of an extended-coil condenser over which air is driven by fans and which is continually bathed in a water spray. Efficient condensing effect is obtained by the combined effect of air cooling and evaporation of water by the hot condenser coils. Small sizes for capacities up to 5 tons use propeller fans and direct tap-water connection for spray. Larger sizes (condensers in excess of 250 tons capacity have been built) are quite similar to unit air conditioners and employ centrifugal fans and a recirculating-spray system with a centrifugal pump.

Maintenance difficulties attendant upon the use of either cooling towers or evaporative condensers have resulted in a vast increase in the application of air-cooled condensers for reciprocating refrigeration systems in capacities up to 200 tons and possibly higher. While the use of air-cooled condensers results in a considerable operating-power penalty under maximum temperature condition, the apparatus employed is simple and relatively maintenance-free, factors which frequently justify the power penalty resulting from their use.

Refrigeration Compressors. Three basic types of compressor are used in refrigeration and electric chiller applications: reciprocating, rotary, and centrifugal.

Reciprocating compressors are typically used in applications below 250 tons. Reciprocating compressors are less efficient than other compressor designs, using between 0.75 and 1.0 kW/ton.

The design of reciprocating compressors has been profoundly influenced by development of the production-line internal-combustion engine as offered today for automotive and aircraft service. Multicylinder in-line arrangements of 2 to 8 cylinders are thoroughly accepted, as are V and W arrangements of 4 to 16 cylinders and radial arrangements of 3, 5, and 7 cylinders. Practically all machines are single-acting with enclosed crankcases. With multicylinder design, large capacities are obtainable with good balance and relatively low piston speed, so that it is not unusual for machines of up to 120 tons capacity, with forced-feed lubrication, to operate at speeds as high as 1750 r/min. Speeds below 500 r/min (except with very large machines) are the exception rather than the rule. Capacity variation is obtained by unloading groups of cylinders or by bypass or clearance-pocket arrangements.

The overall application of refrigerated air conditioning to large industrial and commercial buildings requires capacities far beyond the range of the largest ice-making or storage refrigeration installations. These requirements have been met adequately by development in the *centrifugal refrigeration* system, which finds current application in single units from 200 to 1800 tons or over in capacity. Drive motors will vary from a low of 0.85 hp/ton for industrial applications to a high of 1.05 hp/ton for commercial installations.

Rotary compressors are available in a number of different versions, including vane, scroll, and screw designs. Screw compressors are used in applications up to 1200 tons. Rotary-vane compressors are used in a wide range of applications from small residential refrigeration units to large, rugged industrial processing plants.

Scroll compressors are typically used in applications below 15 tons. Scroll compressors have some key operating advantages in terms of efficiency and reliability. Although scroll-compressor designs were developed in the early 1900s, their complicated geometry discouraged mass production and the design was never fully commercialized until the 1970s. Advances in manufacturing techniques have enabled the cost-effective production of scroll compressors. As a result of certain performance and reliability advantages, the use of scroll compressors is increasing. Scroll compressors are typically used in smaller refrigeration units, less than 15 tons. Screw compressors are used over a wide range of system capacities, from 50 to 1200 tons. Versions of screw compressors include single helical screws and double-helical screws. Screw compressors operate between 0.6 and 0.75 kW/ton.

Centrifugal compressors are commonly used in applications as large as 2000 tons and in systems as small as 80 tons. Centrifugal compressors offer many advantages, including high efficiency and good reliability. Centrifugal units operate between 0.55 and 0.8 kW/ton.

The original designs for centrifugal systems employed multistage compressors with shaft seals and operating at speeds from 3500 to 7000 r/min. Standard 4-pole or 6-pole open motors with step-up gears were customarily used for compressor drives. Wound-rotor motors affording 25% speed regulation were frequently employed for capacity control. However, since 1955 design has largely been directed toward hermetic systems in which the drive motor is enclosed in the gas passage with the compressor, and with this arrangement the drive is naturally limited to the use of the squirrel-cage induction motor. Control of compressor capacity is afforded by regulating dampers in the compressor suction connection. Common designs employ either two-stage compressors direct-driven by 2-pole motors at approximately 3500 r/min or single-stage compressors operating at 11,000 r/min or higher and driven through step-up gears by enclosed 2-pole motors. One design employs an external frequency converter providing 300-Hz current for operation of a 2-pole induction motor for direct compressor drive at approximately 17,500 r/min. Centrifugal compressors are still offered without the enclosed motor and with the shaft extended through the compressor housing with a suitable seal; as such they can employ the open-motor or turbine drive. However, the general design and control arrangement of the system is generally the same for either hermetic or open drives for manufacturing standardization. Even with open units induction-motor drives are generally selected, for though motor sizes are large, average experience indicates that even where power- factor penalties are involved, power-factor correction can be obtained more economically by the use of capacitors in connection with induction motors rather than through the use of synchronous motors. Where closed motors are employed for compressor drive, they are cooled either by water jacketing or by an arrangement employing a bypass of refrigerant or refrigerant gas for motor cooling.

Compressor Drives. While conventional open motors and V-belt drives are generally employed for large-capacity, low-speed reciprocating compressors, which are still offered, the bulk of the production of reciprocating equipment is now of the direct-drive design generally employing 4-pole 1750-r/min motors. This design permits an extremely compact arrangement which is well suited to the unit-system combinations now widely offered by the industry.

In both the reciprocating and centrifugal compressor fields there is contention as to whether the use of standard open motors or sealed hermetic arrangements is most acceptable. The use of standard open-motor drives necessitates the shaft seal, which offers some disadvantages. The use of the hermetic arrangement requires careful design of the electrical components and close control of the refrigerant atmosphere, which must be kept absolutely moisture-free if electrical problems are to be avoided. Obviously, if failure occurs, the repair of the hermetic motor is a far more serious problem than the repair of the open motor. On the other hand, if properly designed and installed the fully enclosed motor operates in a controlled atmosphere and may be less likely to give trouble. However, with either arrangement the motor is generally a standard 2-pole or 4-pole induction motor as far as the basic electrical design is concerned.

Reciprocating systems generally employ means of unloading by holding suction valves open to afford capacity control in operation and are similarly arranged to start unloaded. Centrifugal systems employ dampers for capacity control, which are kept closed during the starting cycle so that these units also start unloaded. However, the characteristics of the reciprocating system are such that the starting-torque requirement is fairly high. Accordingly, line start for drive motors is preferred wherever utility-company regulations will permit, and since these units are usually to approximately 125-hp maximum size and are frequently installed in multiple units of smaller capacity, line start is generally permissible. Where reduced voltage starting is required, autotransformer or resistance-type reduced-voltage starters are generally employed, although frequently these may be effective only as a means of reducing inrush, as in many instances full line voltage must be applied before the compressor will actually pick up the load.

Centrifugal refrigeration systems generally represent larger-capacity units than reciprocating systems, with motor sizes ranging from 75 to 2000 hp or occasionally higher. However, their starting-torque requirement is very moderate, and it is entirely practical to place the compressor in satisfactory operation by use of the Y-Δ starting system, in which the compressor drive motor is connected across the 3-phase source in a Y arrangement during the starting interval and switched to the Δ connection when the compressor has reached practically full speed and before the compressor damper is opened, permitting the unit to assume load. With this arrangement the current during the starting period can be limited to approximately 130% of the normal full-load running current, and this arrangement is generally entirely satisfactory to power companies. This starting arrangement is almost universally employed with centrifugal systems except for relatively small units, which may be within a range suitable for line start.

Motor Voltage. All standard equipment is based on the assumption that 60-Hz current is available, although some manufacturers will still list a capacity rating of their units of 60-Hz design when employed on 50-Hz service. All systems involving motors up through 5 hp are generally designed for 230-V 60-Hz 1-phase service, but in sizes 3 hp and above are equally available for 208/220/440-V 60-Hz 3-phase service, which is generally available in most urban areas. Equipment for practically the entire available capacity range is designed for service on 208-V 60-Hz 3-phase current, which is commonly available as urban low-voltage distribution. However, it should be noted that because of the rigorous nature of compressor service it is usually inadvisable to attempt to use 220-V equipment for the compressor drive where only 208 V is supplied. Many feel that good design dictates the use of 200-V motors for such service. For larger commercial or industrial applications which are provided with primary service through independent substations, 440-V 60-Hz 3-phase equipment is generally preferred, as it permits a substantial reduction in wiring cost. Currently a number of systems employ 480/277-V 60-Hz 1-phase 4-wire service, where the higher single-phase voltage may be satisfactorily used for fluorescent-lamp operation. This arrangement affords considerable economy in circuit design but presents complications with respect to the use of standard 115-V 60-Hz 1-phase conventional power or office equipment, for which separate transformers and circuiting must be provided. In the textile industry extensive use is made of 550-V 60-Hz 3-phase service, and equipment is generally available throughout the entire capacity range for this voltage characteristic. In larger plant operations employing plant primary service at either 2300 or 4160 V, these voltages are frequently used for large compressor drives, and most lines make equipment available for this voltage in the range from 200 to 2000 hp, although economics would generally indicate 500-hp drives as being the lowest acceptable size for this voltage.

Reverse-Cycle or Heat Pump. The reverse-refrigeration cycle is for the purpose of heating interiors with electricity. Heat absorbed at a low-temperature level is pumped to a level sufficient to permit satisfactory heating with the expenditure of only the amount of energy necessary to perform the pumping work, which is also reclaimed at the high-temperature level. Under certain circumstances this cycle is entirely practical. Many installations have been made where well water at moderate temperatures (±50°F) is available as a low-temperature heat source, and these installations have shown coefficients of performance up to 400%, that is, 5 Btu of heating effect for 1 Btu of energy applied. However, the performance of the refrigeration machine is sharply limited as the temperature head between the evaporator and the condenser is increased, so that the practical possibilities of this cycle are affected not only by the cost of electrical energy compared with other fuels but by the level of the low-temperature heat source and the temperature that must be maintained within the interior.

Application of air conditioning to residences has accelerated activity in the heat-pump field. Since the basic refrigeration equipment is being installed for summer conditioning, its utility is increased by arranging it to handle at least a part of the winter heating requirements. Most new developments are directed toward the use of outside air as a low-temperature heat source as well as a means for providing condensing effect when the system is on the cooling cycle. Heating capacity of such a system will drop fairly rapidly as outside temperatures drop. However, there is a balance point between the requirement for summer cooling and available capacity for winter heating which will permit a rationally designed apparatus to care for the full heating requirements on the reverse-cycle basis for substantial periods during the heating season. Naturally, the more temperate the climate in which the apparatus is employed, the greater proportion of the total winter heat requirement can be handled entirely by the heat-pump principle. Since peak heating demand is experienced only for short periods, usually during early morning hours, the current procedure is to supplement the available output of the heat-pump device by direct resistance heating. This represents a substantial spot load, but its period of use is relatively short, and under most utility rate schedules the overall economies of the operation are reasonable. However, it presents problems in power distribution. A moderate-sized residence in the South Atlantic area with a basic 6-kW requirement for summer or heat-pump operation may be supplemented with as much as 12 kW of additional resistance heating load for use during peak heating periods. This imposes a maximum load of up to 18 kW per residential unit on the lines of the utility, with little possible diversity.

On the other hand, this load occurs at a time when other urban loads are at a minimum and when outside ambient-temperature conditions are at a very low level, so that distribution circuits, transformers, etc., can handle a substantial overload with no difficulty. Accordingly, utility companies are finding that they can handle this load with no serious problems and feel that it is attractive.

A number of ingenious arrangements employing the heat-pump principle or other heat-recovery systems have been incorporated in designs of installations for large commercial buildings. It is evident that the core sections of these buildings are subject to cooling requirements throughout the entire year, as they are shielded from exterior exposure, but subject to heat gains from lighting and occupancy. A number of systems have been employed which use refrigeration for these interior zones throughout the entire year but during the wintertime take the effluent heat from the condenser of the refrigeration system and apply this as a means of heating the periphery of the building which is subject to normal winter exposure. Thus, in effect, the excess heat from the interior core of the building is transferred to the exterior for heating purposes while affording control of the heat load within the core of the building.

Other arrangements provide for direct cooling of lighting equipment either by circulating water or by exhaust ventilation and transfer the heat thus directly recovered to the periphery of the building for heating purposes. There are obviously many possibilities of this nature which must be studied for each individual project. However, where this means of heating is used, it is dependent on a continued supply of electrical energy to the space as a primary heat source. In some instances arrangement is made to turn on interior lighting during cold periods for the purpose of heating the premises, or where lights are turned off, the equivalent amount of electrical energy may be applied through resistance heating, thus affording a heat source with no increase in electrical demand. These arrangements frequently make full electric heating entirely practical and economical for many large buildings.

Refrigeration Storage: Moderate Temperature. Many refrigeration applications having high peak loads for short periods of time can be most economically handled by the use of relatively small refrigeration machines which are operated over long periods of time to store refrigeration effect to meet the high short-period requirement. Refrigeration effect may be stored in bodies of brine or water, though a relatively large storage volume will be required, since the water or brine can only absorb or give up its sensible heat. A more efficient storage method involves the formation of ice on submerged evaporating coils. Every pound of ice thus formed absorbs and can release, on melting at 32°F, its latent heat of fusion, 144 Btu/lb.

This principle has been frequently used to improve the load factor of refrigeration systems used in conjunction with air-conditioning systems having relatively short use periods (restaurants serving only one meal a day, theaters with only evening showings, and churches or infrequently used auditoriums).

Refrigeration Storage: Low Temperature. A solution of water and salt (or, in general, of any two substances) has a certain concentration which results in the lowest freezing temperature. A solution of this concentration is known as a *eutectic mixture*. A brine with a salt content lower than this concentration will start to freeze above this minimum temperature, and as cooling progresses, pure ice crystals will freeze out, thus increasing the concentration of the brine until the eutectic concentration is reached. At this point freezing will take place at a constant temperature until all latent heat is removed, the resulting crystals being a mechanical mixture of salt and frozen water. If the starting salt concentration of the mixture is greater than the eutectic concentration, salt crystals will first freeze out until the eutectic concentration is reached.

The use of eutectic mixture as a means of storing refrigerating effect, that is, latent heat of fusion, at a temperature below 32°F (the freezing point of water ice) has received wide attention. The eutectic mixture of water 76.7% and sodium chloride (common salt) 23.3% has a freezing temperature of –6°F, and a latent heat of fusion of 101.5 Btu/lb is available at this temperature. Plate evaporators charged with eutectic mixtures are frequently used in refrigerator trucks which are cooled at night to produce sufficient stored cooling effect to meet the next day's operation.

BIBLIOGRAPHY

Sources of additional information on the use of electricity in agriculture are as follows:

Agricultural experiment stations and cooperative extension services, state colleges.

Farm Electrification Section, Agricultural Research Service, U.S. Department of Agriculture (USDA), Beltsville, Md.

National Rural Electric Cooperative Association, 1800 Massachusetts Avenue N.W., Washington, D.C. 20036.

Richey, C. B., Jacobson, P., and Hall C. W., *Agricultural Engineers' Handbook*; New York, McGraw-Hill Book Company, 1961.

Rural Utilities Services, USDA, Washington, D.C.

SECTION 22
POWER ELECTRONICS

Amit Kumar Jain
Engineering Technical Staff, Analog Power Design Inc.

Raja Ayyanar
Associate Professor, Department of Electrical Engineering, Arizona State University

CONTENTS

22.1 INTRODUCTION

22.1.1 Role of Power Electronic Converters

Power electronics is an enabling technology that achieves conversion of electric power from one form to another, using a combination of high-power semiconductor devices and passive components— chiefly transformers, inductors, and capacitors. The input and output may be alternating current (ac) or direct current (dc) and may differ in magnitude and frequency. The conversion sometimes involves multiple stages with two or more converters connected in a cascade. The end goals of a power electronic converter are to achieve high efficiency of conversion, minimize size and weight, and achieve desired regulation of the output. Figure 22-1 shows power electronic converters in a generic application.

22.1.2 Application Examples

Power electronic converters can be classified into four different types on the basis of input and output, *dc-dc, dc-ac, ac-dc,* and *ac-ac,* named with the first part referring to the input and the second to the output. The diode bridge rectifier is the front end for most low-power converters. It converts line frequency ac (e.g., from a wall outlet) to an unregulated dc voltage, and the process is commonly called rectification. In a *dc-dc* converter, both the input and the output are dc, and in the simplest case the output voltage needs to be regulated in presence of variation in load current and changes in the input voltage. A computer power supply has a diode bridge front end followed by a dc-dc converter, the combination of which converts line frequency ac voltage to several regulated dc voltages (Fig. 22-2). Electronic ballasts for compact fluorescent lamps consist of a line frequency rectifier followed by a dc to high-frequency ac converter (frequency range of 20 to 100 kHz) whose output is connected to a resonant tank circuit that includes the load. In an adjustable speed motor drive application (Fig. 22-3), the input is a 3-phase ac supply, and the output is a 3-phase ac whose magnitude and frequency are varied for optimum steady-state operation and dynamic requirements of the drive.

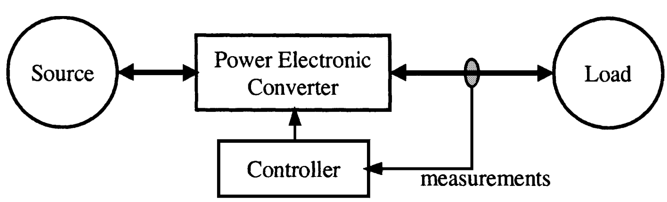


FIGURE 22-1 Application of power electronic converters.

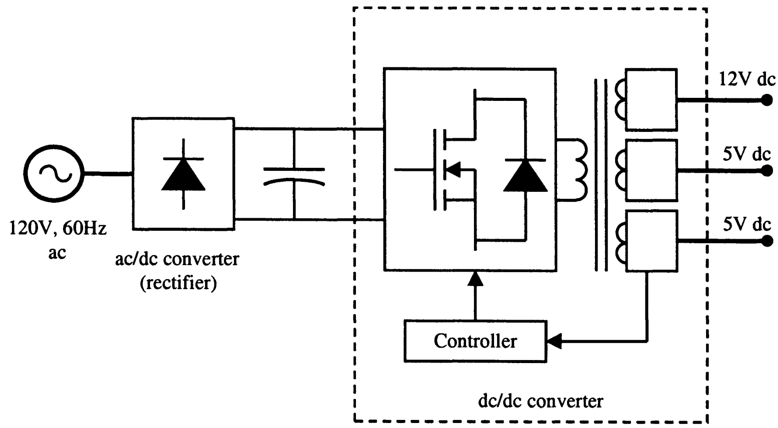


FIGURE 22-2 Computer power supply.

Development of power semiconductors with very high voltage and current ratings has enabled the use of power electronic converters for utility applications. In transmission systems, power electronic converters are being utilized to control power flow, damp power oscillations, and enhance system stability. At the distribution level, power electronic converters are used for enhancing power quality by means of dynamic voltage restorers, static var compensators, and active filters. Power electronic converters also play a significant role in grid connection of distributed generation and especially renewable energy sources; their functions include compensation for steady state and dynamic source characteristics leading to optimal energy transfer from the source, and protective action during contingencies.

Future automotives are expected to have a large number of power electronic converters performing various functions, for example, electric power steering, active suspension, control over various loads, and transferring power between the conventional 14-V bus and the recently proposed *42-V Power Net* [1]. Hybrid electric and all-electric vehicles also utilize controlled power electronic converters for interfacing the battery and motor/generator.

The proliferation of power electronics connected to the utility grid has also led to power quality concerns due to injection of harmonic currents by grid-connected inverters, and highly distorted currents drawn by diode bridge rectifiers. Due to fast transients of voltages and currents within power

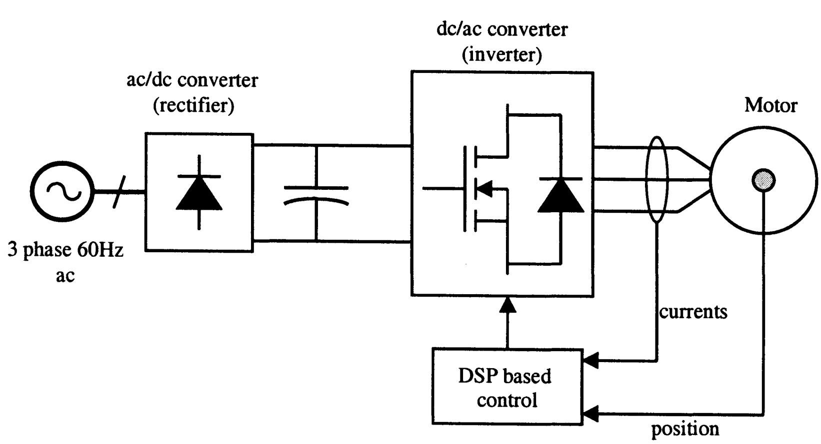


FIGURE 22-3 Adjustable speed motor drive.

converters, they can be a source of electromagnetic emissions leading to electromagnetic interference. Several solutions to limit and correct these effects have therefore been developed.

22.1.3 Scope and Organization

This section gives an overview of power electronic systems. Details of specific converter types and applications have been omitted and only the fundamentals are presented. In some cases, important results are stated without derivation. Mathematical content has been kept to a minimum. In places, empirical aspects have been included, since power electronics is an application-oriented discipline. Design procedures are presented with only those justifications that were deemed imperative. A long list of references consisting of textbooks on the subject of power electronics, reference books on specific areas and applications of power electronics, important research publications, and several online sources has been provided. The reader is expected to use this section as a starting point, followed by the references on the topic of particular interest.

First, the basic principles for analysis and design of power converters are presented in Sec. 22.2. Topology and operating principles of the four types of power electronics converters are described with one section devoted to each. A very simple description of power electronic converter control is presented using the example of dc-dc converters. This is followed by deleterious effects of power electronic converters and precautions necessary to limit or correct them. Applications are described next bringing together the requirements and complete power electronic system realization for some specific examples. Finally, the individual components that constitute a power electronic converter are discussed. Current research initiatives and expected future trends are indicated in each section.

22.2 PRINCIPLES OF SWITCHED MODE POWER CONVERSION

This section presents some basic principles that are common to the analysis of all switch mode power converters. Line-commutated power electronic converters are not, strictly speaking, switched mode converters; they are discussed in Sec. 22.6.

22.2.1 Bipositional Switch

The most basic component of a switch mode power converter is the bipositional switch shown in Fig. 22-4*a*. Nodes 1 and 2 of the switch are invariably connected across a dc voltage source (or across

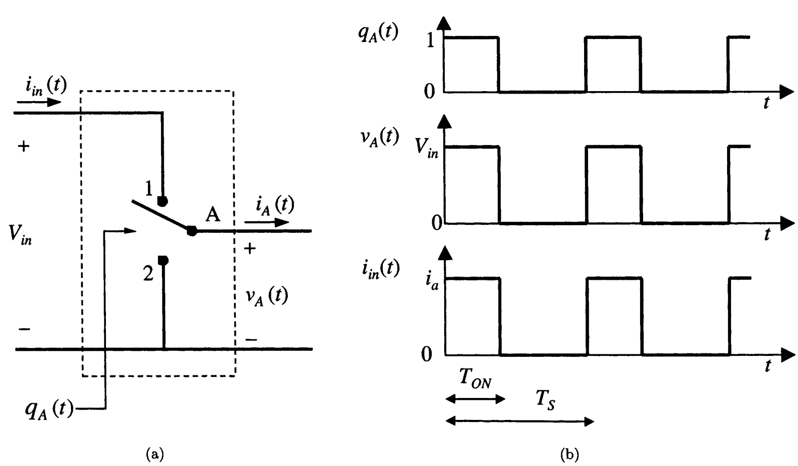


FIGURE 22-4 (*a*) Bipositional switch (*b*) switching waveforms.

TABLE 22-1 States of a Bipositional Switch

$q_A(t)$	Switch position	MOSFET & diode state	Pole voltage & input current
1	1	S1 & D1 ON, S2 & D2 OFF	$v_A = V_{in}$, $i_{in} = i_A$
0	2	S1 & D1 OFF, S2 & D2 ON	$v_A = 0$, $i_{in} = 0$

a big capacitor whose voltage is close to a constant dc), and pole "A" of the switch is in series with a dc current source (or a big inductor whose current is close to a constant dc). This bipositional switch, which is also referred to as a switching power pole, switches at very high frequencies, and is controlled by the signal $q_A(t)$. The switched pole A voltage and the input current based on the control signal $q_A(t)$ are listed in Table 22-1, and the corresponding waveforms are shown in Fig. 22-4*b*.

Figure 22-5*a* shows the electronic implementation of a complete bipositional switch using metal-oxide-semiconductor field-effect transistors (MOSFETs). This implementation can support pole current in either direction and is useful for applications where current direction can reverse. In most dc-dc converter applications, the current through the pole *A* is unidirectional, and hence, the implementation shown in Fig. 22-5*b* is sufficient to realize the bipositional switch.

22.2.2 Pulse Width Modulation

The concept of pulse width modulation (PWM) is central to all switch mode power converters. Pulse width modulation refers to the control of the average value of a switching variable, for example, $v_A(t)$ in Fig. 22-4*b*, by controlling or modulating its pulse width. Some basic concepts and definitions necessary to understanding PWM are presented here.

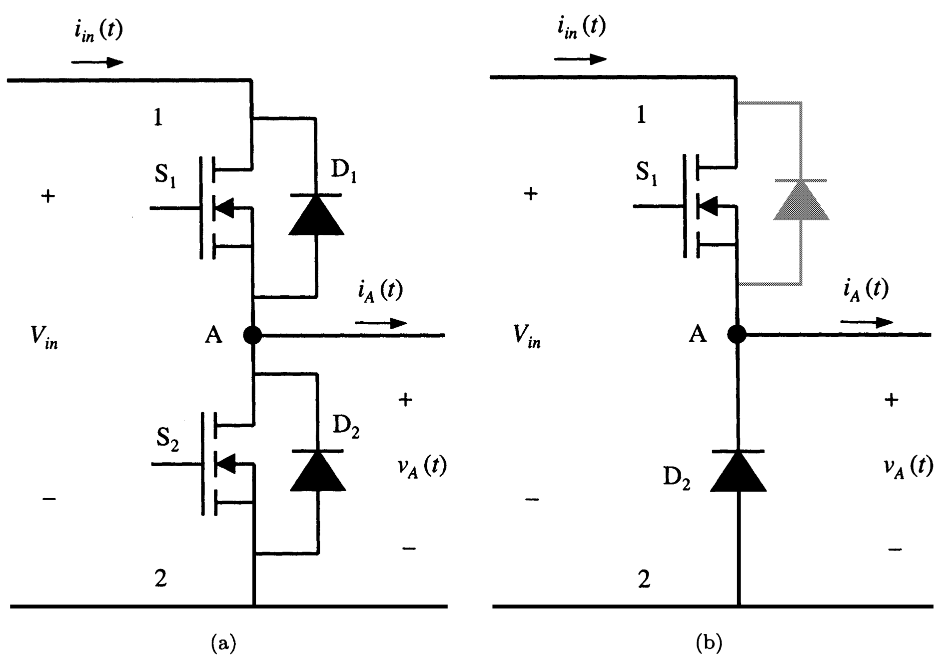


FIGURE 22-5 Electronic implementation of bipositional switch: (*a*) for bidirectional pole current (*b*) for unidirectional pole current.

Duty Ratio. The frequency at which the bipositional switch is switched on and off is denoted by f_s, and the corresponding time period by T_s (= $1/f_s$). The transition between the two states of the switch occurs in a very small duration compared to T_s. The time for which the switch remains in position 1 during a switching period is denoted by T_{on}. The duty ratio d of the bipositional switch is then defined as the ratio of on-time to total time period:

$$d = \frac{T_{on}}{T_s} \tag{22-1}$$

Averaging. Currents and voltages in power electronic converters have (1) high-frequency components corresponding to the switching frequency of the bipositional switch elements, and (2) low-frequency components due to slower variations caused by change in load demand, source magnitude, and changes in reference value of the desired outputs. For dynamic control and steady-state analysis, the low-frequency components are of primary interest. To study these components, it is sufficient to study their averages over one switching time period. It should be noted that the averaging presented here [2] is a very basic form of the general averaging method [3] and has limitations in terms of validity with respect to the switching frequency. However, this simplification is good enough for most practical purposes, and can be confidently used for steady state and dynamics up to one-fifth the switching frequency. Throughout this chapter the averaged variables, that is, averaged over one switching period, are denoted by a "-" on top of the variables. Thus, the averaged value of $x(t)$ is given by

$$\bar{x}(t) = \frac{1}{T_s}\int_{t-T_s}^{t} x(\tau)\, d\tau \tag{22-2}$$

In steady state, the average values of $q_A(t)$ and $\upsilon_A(t)$ are given by

$$\bar{q}_A = \frac{1}{T_s}\int_0^{T_s} q_A(\tau)\, d\tau = \frac{1}{T_s}\int_0^{T_{on}} 1\, d\tau = \frac{T_{on}}{T_s} = d \tag{22-3}$$

$$\bar{\upsilon}_A = \frac{1}{T_s}\int_0^{T_s} \upsilon_A(\tau)\, d\tau = \frac{1}{T_s}\int_0^{T_{on}} V_{in}\, d\tau = d \cdot V_{in} \tag{22-4}$$

In general, the averaged quantities can be time varying, since the pulse widths of the switching waveform can vary with time. Thus

$$\bar{q}_A(t) = d(t) \tag{22-5}$$

$$\bar{\upsilon}_A(t) = d(t) \cdot V_{in} \tag{22-6}$$

As an example of PWM, we can regulate the average value of $\upsilon_A(t)$ in Fig. 22-4*b* by varying the duty ratio d. If $V_{in} = 10$ V, $f_s = 100$ kHz $\Rightarrow T_s = 10\ \mu$s, then $T_{on} = 5\ \mu$s $\Rightarrow d = 0.5$, and $\upsilon_A = 5$ V, etc. By varying the duty ratio sinusoidally a low-frequency ac voltage can be synthesized from a dc voltage, as illustrated in Fig. 22-6.

22.2.3 Concept of Steady State

A converter is said to be in dc steady state when all its waveforms exactly repeat in each switching period, that is, $x(t) = x(t - T_s)\ \forall\ t$, where x is any of the converter variables. With reference to Eq. (22-2), it is clear that in steady state the average value of any variable is constant. Analysis of steady-state operation is essential to determine ratings and design of the power stage components in the converter, viz, power semiconductor devices, inductor, capacitors, and transformers. Important

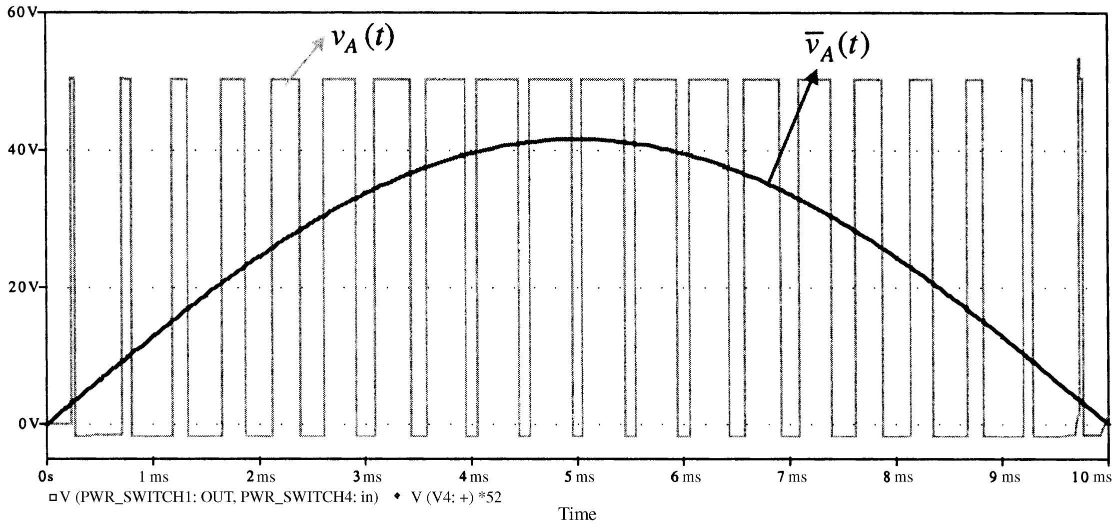


FIGURE 22-6 AC synthesis using PWM.

concepts that enable steady-state analysis from a circuit view point are discussed below. It should be remembered that these are only valid during steady-state operation.

Steady-State Averages of Inductor Voltage and Capacitor Current. The instantaneous v-i relationship for an inductor is

$$v_L(t) = L\frac{di_L(t)}{dt} \quad \text{or} \quad i_L(t) = i_L(0) + \frac{1}{L}\int_0^t v_L(\tau)\,d\tau \tag{22-7}$$

where $v_L(t)$ is the voltage across the inductor and $i_L(t)$ is the current flowing through the inductor. Since $i_L(T_s) = i_L(0)$ in steady state, from the integral form of Eq. (22-7) it follows that

$$\bar{v}_L = \frac{1}{T_s}\int_0^{T_s} v_L(\tau)\,d\tau = 0 \tag{22-8}$$

The above relationship can also be derived directly in terms of the averaged quantities as follows:

$$\bar{v}_L(t) = L\frac{d\bar{i}_L(t)}{dt} = 0 \quad \text{(since } \bar{i}_L(t) \text{ is constant in steady state)} \tag{22-9}$$

This is referred to as *volt-second balance* in an inductor. Figure 22-7 shows a typical steady-state waveform of an inductor voltage for many power converters. The positive area is exactly cancelled by the negative area, making the average value zero. It may be mentioned that during the start-up transient, $\bar{v}_L$ remains positive for several switching cycles, allowing the inductor current to rise from zero to its final steady-state value.

In a similar fashion, it can be shown that in steady state the average current through a capacitor

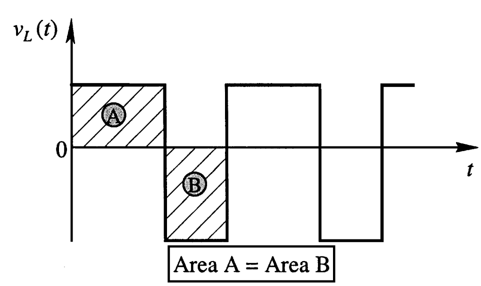


FIGURE 22-7 Volt-second balance for an inductor.

is zero. This is referred to as *ampere-second balance* in a capacitor. Note that though the average value of the capacitor current is zero, its root mean square (RMS) value, which is one of the main selection criteria for a capacitor, can be substantial depending on the converter topology.

Power Balance. For analytical purposes, it is often useful to neglect all losses in the converter and consider input power to be equal to the output power, again in an average sense

$$P_{in} = \bar{v}_{in}\bar{i}_{in} = P_o = \bar{v}_o\bar{i}_o \tag{22-10}$$

This implies that there is no increase or decrease in the energy stored in inductors and capacitors over one switching time period. This is valid for the input-output of the entire converter as well as any intermediate stage.

Kirchoff's Laws for Averages. Just like the instantaneous quantities, the averaged quantities also obey Kirchhoff's current and voltage laws. The sum of average currents entering a node is zero. The proof follows from interchanging the order of summation (for individual currents) and integration (over a switching time period)

$$\sum_k \bar{i}_k = \frac{1}{T_s}\sum_k\left[\int_0^{T_s} i_k\right] = \frac{1}{T_s}\int_0^{T_s}\left[\sum_k i_k\right] = 0 \qquad \left(\text{since } \sum_k i_k \equiv 0\right) \tag{22-11}$$

Similarly, the sum of average voltages in a circuit loop is zero.

$$\sum_k \bar{v}_k = 0 \tag{22-12}$$

22.2.4 Power Loss in the Bipositional Switch

Electronic implementations of the bipositional switch shown in Figs. 22-5*a* and 22-5*b* have significant power loss. The power loss can be divided into two kinds—conduction loss and switching loss.

With reference to Fig. 22-5*b,* when the MOSFET is on there is a nonzero voltage across it. Similarly the diode has a forward voltage drop while it is conducting. Both of these lead to power loss whose sum averaged over one switching time period is called conduction loss.

A finite time interval is required to transition from one state to the other: (MOSFET on and diode off) to (MOSFET off and diode on), and vice versa. While the MOSFET is turning off, the diode cannot conduct until it is forward biased. As the voltage across the MOSFET increases from near zero to the full input voltage V_{in}, it conducts the full output current. Once the diode is forward biased the current starts transferring from the MOSFET to the diode. During the reverse transition, first current is transferred from the diode to the MOSFET, and then the voltage across the MOSFET reduces from V_{in} to the conduction voltage drop. Thus, the MOSFET incurs significant power loss during both transitions. The above description is simplified and there are other phenomena which contribute to loss during the transitions. The diode also has power loss during the transitions. The sum of losses in the MOSFET and diode during the transitions averaged over one switching time period is called the switching loss. Switching power loss increases with increase in switching frequency and increase in transition times. Sum of the conduction and switching loss, computed as averages over one complete switching periods, gives the total power loss Fig. 22-8.

Similar losses occur in the realization of Fig. 22-5*a*. When S_1 is turned off by its control signal, current $i_A(t)$ transfers to D_2, the antiparallel diode of S_2. After this transition, S_2 is turned on and the current transfers from the diode to the MOSFET channel (which can conduct in either direction). A short time delay, called dead time, is required between the on signals for S_1 and S_2. The dead time prevents potential shorting of the input voltage, also known as shoot-through fault.

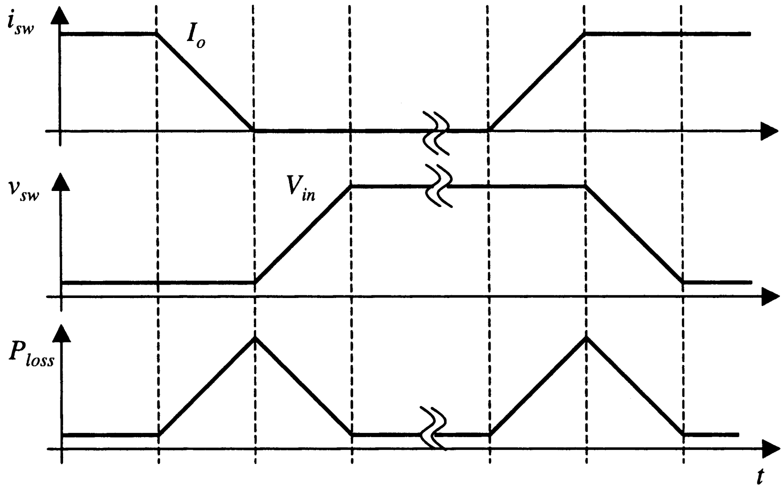


FIGURE 22-8 Switching transients in bipositional switch implementation.

The nonidealities of nonzero voltage drop and switching times will be neglected for analysis of power electronic converters presented throughout this chapter. However, these are extremely important in design and selection of components for a real power converter.

22.3 DC-DC CONVERTERS

DC-DC converters represent one major area in power electronics. In a dc-dc converter, the input and output may differ in magnitude, the output may be electrically isolated from the input, and the output voltage may have to be regulated in the presence of variation in input voltage and load current. In a typical power distribution system (for digital systems), several lower magnitude dc voltages are derived from a common input using a one or more converters. Battery-powered portable devices use converters that boost the input 1.5 V cell voltage to 5 or 9 V. Most of these converters have unidirectional power flow—from input to output. The presentation here is limited to the basic converter types. The interested reader is referred to text books that deal with details of these converters [4–8].

22.3.1 Buck Converter

The buck converter is used to step down an input voltage to a lower magnitude output voltage. Figure 22-9*a* shows the schematic of a buck converter. A power MOSFET and diode combination is shown for implementation of the bipositional switch with unidirectional output current. The bipositional switch is followed by an L-C low-pass filter that attenuates the high-frequency switching component of the pole A voltage and provides a filtered dc voltage at the output. A high switching frequency is desirable to reduce the size of the filter, the higher limit depending on the power level of the converter and the semiconductor devices used. The final choice of switching frequency depends on several factors: size, weight, efficiency, and cost. It is usually above the audible range and frequencies above 100 kHz are very common.

Operation. The input voltage V_{in} is assumed to remain constant within a switching cycle. The inductor L and capacitor C are sufficiently large so that the inductor current i_L and output voltage v_o do not change significantly within one switching cycle. The load is represented by the resistor R_L. Under

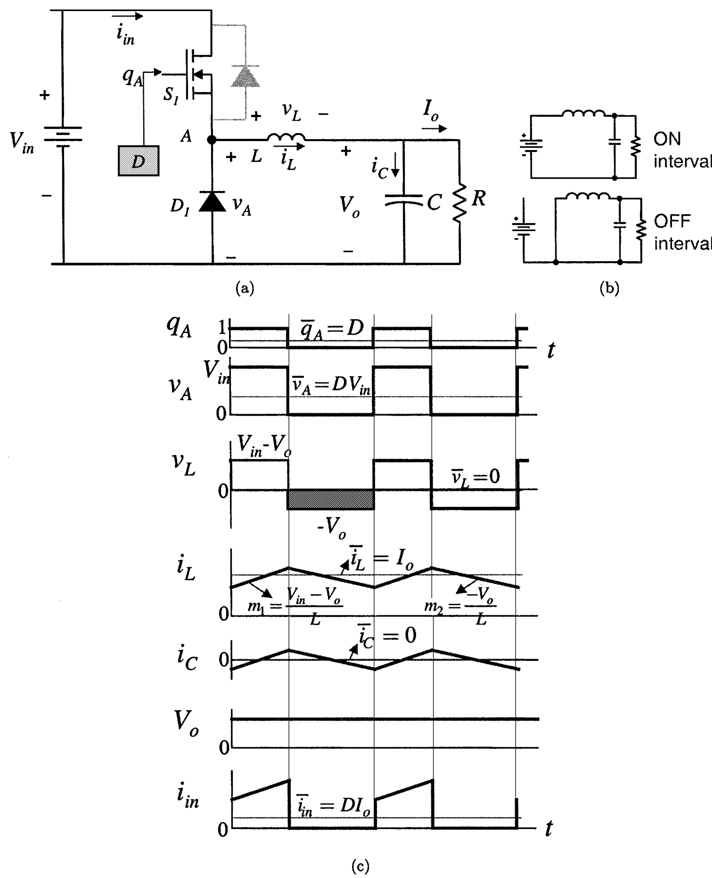


FIGURE 22-9 Buck converter: (*a*) circuit, (*b*) equivalent circuits during ON and OFF intervals, (*c*) steady-state waveforms.

steady-state operation it is assumed that the inductor current is always greater than zero. The MOSFET is turned on in response to signal $q_A(t)$ for $T_{on} = DT_s$, where D represents the steady-state duty ratio. During this time $v_A = V_{in}$ and $i_{in} = i_L$. When the MOSFET is turned off, the inductor current flows through diode D_1 leading to $v_a = 0$ and $i_{in} = 0$. Since the average voltage across the inductor is zero, $\bar{v}_L = 0$, the average output voltage is given by

$$\bar{v}_o = \bar{v}_a = DV_{in} \tag{22-13}$$

The average current through the capacitor C, $\bar{i}_c$, is zero. Thus, $\bar{i}_L = I_o$ and the input current is given by

$$\bar{i}_{in} = DI_o \tag{22-14}$$

From the above equations it is clear that the output voltage is lower than the input voltage and output current is higher than the input current. Also, power balance for averaged quantities can be verified from Eqs. (22-13) and (22-14). Within a switching cycle, instantaneous values of the inductor current and capacitor voltage vary as follows:

$$\begin{array}{lll} \text{MOSFET on:} & L\dot{i}_L = V_{in} - v_o & C\dot{v}_o = i_L - v_o/R_L \\ \text{MOSFET off:} & L\dot{i}_L = -v_o & C\dot{v}_o = i_L - v_o/R_L \end{array} \tag{22-15}$$

Equivalent circuits for the two intervals and instantaneous waveforms are shown in Figs. 22-9*b* and 22-9*c*.

Component Selection. Usually the inductor and capacitor are significantly large so that within a switching period v_o can be assumed constant in computation of i_L. This leads to the linear variation of i_L shown in Fig. 22-9*c* with a *peak-to-peak ripple* ΔI_L given by

$$\Delta I_L = \frac{V_o(1 - D)T_s}{L} \tag{22-16}$$

In most designs, the inductance value is chosen to limit ΔI_L between 10% and 30% of the full load current. Since the average capacitor current is zero, the instantaneous capacitor current is approximately equal to the ripple component of the inductor current.

$$i_C(t) = i_L(t) - I_o \tag{22-17}$$

The peak-to-peak capacitor voltage ripple resulting from the capacitor current can then be derived as

$$\Delta V_o = \frac{\Delta I_L \cdot T_s}{8C} \tag{22-18}$$

Capacitors used for filtering in most dc-dc converters are electrolytic capacitors, which are characterized by significant effective series resistance (ESR) and effective series inductance (ESL). These parasitics also contribute to the output voltage ripple and should supplement Eq. (22-18) in the choice of capacitors. Film or ceramic capacitors, which have significantly lower ESR and ESL, should be used in conjunction with electrolytic capacitors.

The MOSFET has to be rated to block a voltage greater than V_{in}, and conduct an average current greater than I_{in}. Power dissipation and temperature considerations usually require MOSFETs to be rated from 2 to 3 times the maximum input average current. In addition, the peak MOSFET current, equal to the maximum peak of the inductor current, should not exceed its maximum current rating. The diode has to be rated to block V_{in}, and conduct an average current greater than the maximum output current. Diodes are usually chosen with ratings approximately 2 times the expected maximum current.

PWM Control Implementation. As evident from Eq. (22-13) the duty ratio of the switch controls the output voltage. In response to variation in input voltage and load current, the duty ratio has to be changed by a feedback controlled system as shown in Fig. 22-10*a*. The error between the reference and actual output voltage is given to an appropriately designed error compensating amplifier, the output of which is a control voltage v_c. This control voltage is compared with a constant frequency sawtooth waveform. The output of the comparator is the switching signal $q_A(t)$ that determines the on or off state of the MOSFET. When the output voltage is lower than the reference value, the control voltage increases, leading to an increase in the duty ratio, which in turn increases the output voltage. The error amplifier and comparator, and several other features, are available in a single

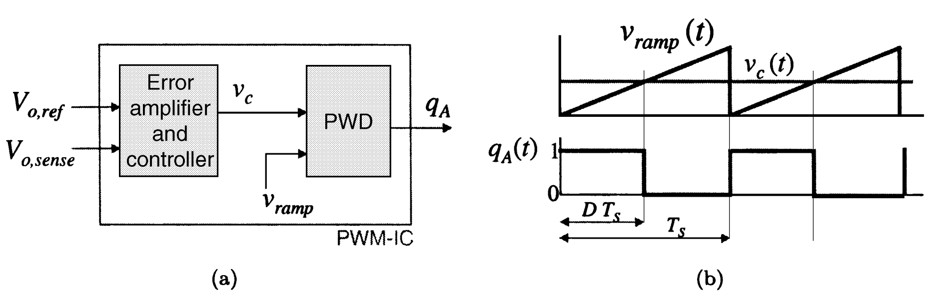


FIGURE 22-10 PWM generation: (*a*) ramp comparison, (*b*) control block diagram.

integrated circuit (e.g., UC3825A) available from several manufacturers (e.g., see Refs. [9–12]). This integration of components leads to reduction in overall size and cost.

22.3.2 Boost Converter

As evident from the name, the boost converter is used to step up an input voltage to a higher magnitude output voltage see Fig. 22-11*a*. In this case, the MOSFET is in the lower position while the diode is in the upper position. The inductor is on the input side and the output has a purely capacitive filter.

Assumptions made for analysis of buck converter are made here as well. When the MOSFET is on in response to $q_A(t) = 1$, diode D_1 is off, and the inductor current increases due to a positive voltage across it (Fig. 22-11*b*). When the MOSFET is switched off, the inductor current flows through diode D_1 and its magnitude decreases as energy is transferred from the inductor to the output capacitor and load. Instantaneous values of the variables during on and off intervals are

$$\begin{array}{lllll} \text{MOSFET on} & v_A = 0 & i_d = 0 & L\dot{i}_L = V_{in} & C\dot{v}_o = -v_o/R_L \\ \text{MOSFET off} & v_A = v_o & i_d = i_L & L\dot{i}_L = V_{in} - v_o & C\dot{v}_o = i_L - v_o/R_L \end{array} \quad (22\text{-}19)$$

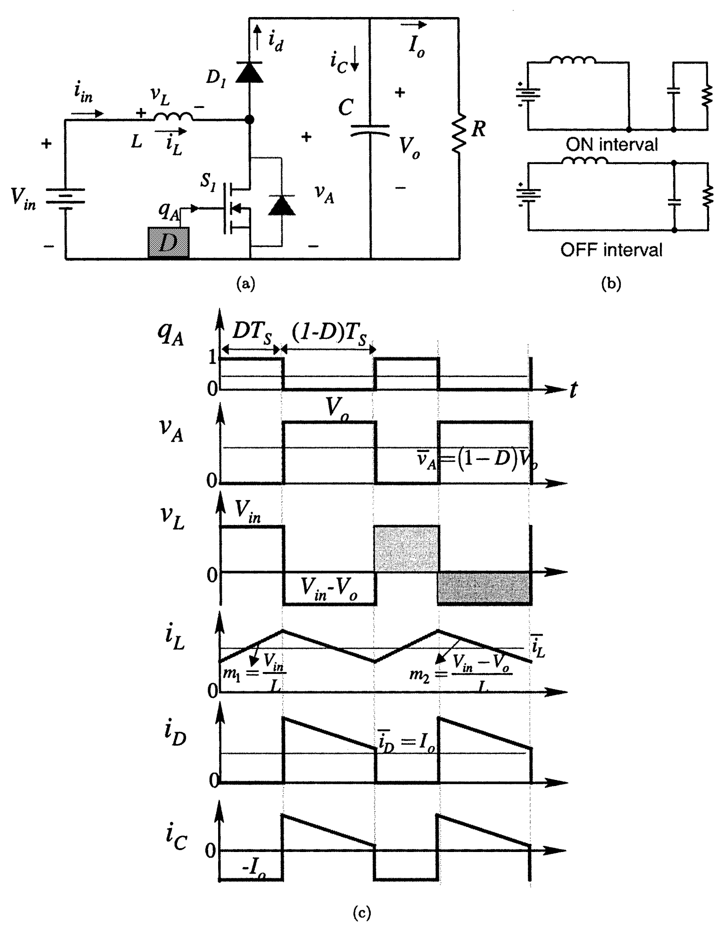


FIGURE 22-11 Boost converter: (*a*) circuit, (*b*) operating states, (*c*) steady-state waveforms.

Noting that $\bar{v}_L = 0$, the averaged pole A voltage is $\bar{v}_A = V_{in} = (1 - D)V_o$. In addition, using $\bar{i}_C = 0$, the steady-state conversion ratios for the boost converter can be obtained as follows:

$$\frac{V_o}{V_{in}} = \frac{1}{1 - D} \qquad \text{and} \qquad \frac{I_{in}}{I_o} = \frac{1}{1 - D} \tag{22-20}$$

From the above equation it is evident that the output voltage is always higher than the input voltage, and conversely, the output current is always lower than the input current by the same ratio. Waveforms of the boost converter variables are shown in Fig. 22-11*c*. The PWM implementation is the same as in the buck converter (Fig. 22-10), with the control objective being regulation of output voltage to a desired value.

The buck and boost converters are capable of either decreasing or increasing the input voltage magnitude, but not both. The buck-boost converter is the third basic converter that can be used to obtain an output voltage both lower and higher than the input voltage; since the output voltage is usually maintained constant, this implies that the input voltage may be higher or lower than the output voltage. A drawback of the buck-boost converter is that the output voltage polarity is inverted with respect to the input voltage return. The SEPIC converter (single-ended primary inductor converter) provides buck and boost gain without polarity inversion but at the expense of additional components. The Ćuk converter, derived from the buck-boost converter using the duality of current and voltage, is another basic dc-dc converter topology [4, 5]. None of the above converters have electrical isolation between the input and the output; however, isolated versions for all of these can be derived.

22.3.3 Flyback Converter

Figure 22-12*a* shows the buck-boost converter circuit. Discussion of this converter in its original configuration is omitted here. Instead, its electrically isolated version known as the flyback converter is described. The flyback converter is very common for low power applications. It has the advantage of providing electrical isolation with low component count. Derivation of the flyback converter from the buck-boost converter is shown in Fig. 22-12*a*. The flyback converter has a coupled inductor instead of an inductor with just one winding. The primary winding is connected to the input while

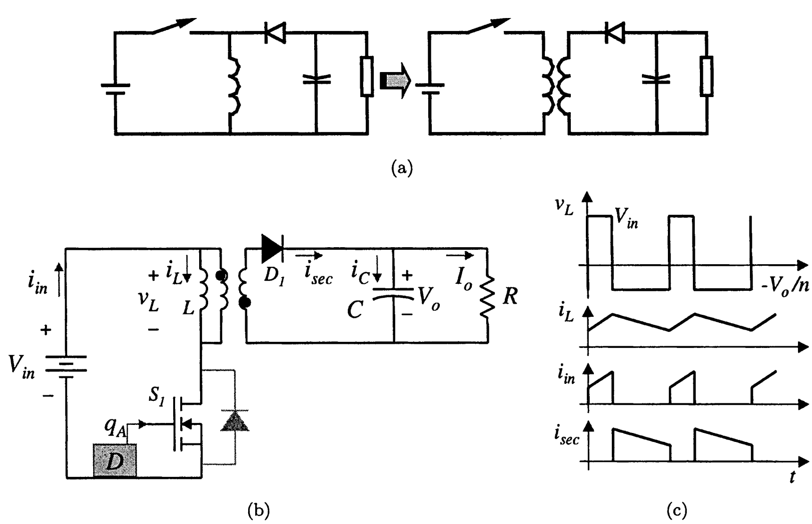


FIGURE 22-12 Flyback converter: (*a*) derivation from buck-boost, (*b*) circuit, (*c*) steady-state waveforms.

the secondary is connected to the output. The circuit diagram is shown in Fig. 22-12*b*. The coupled inductor is represented by an inductor on the primary and an ideal transformer between the primary and secondary. The coupling coefficient is assumed to be 1. Assuming the primary to secondary turns ratio is 1 : *n*, we get

$$\bar{v}_L = V_{in} DT_s - \frac{V_o}{n}(1 - D)T_s = 0 \tag{22-21}$$

$$\Rightarrow \frac{V_o}{V_{in}} = n\frac{D}{1 - D} \tag{22-22}$$

$$\Rightarrow \frac{I_{in}}{I_o} = n\frac{D}{1 - D} \tag{22-23}$$

Although the analysis presented here assumes that the inductor current never goes to zero (called *continuous conduction mode* or CCM), it is very common to design flyback converters so that the inductor current does go to zero within each switching cycle. This operation, known as *discontinuous conduction mode* (DCM), leads to simplification of control design for flyback converters [5]. It should be noted that requirement of electrical isolation is not the only reason that a transformer or (coupled inductor) is used. Another important reason is that the transformer turns ratio leads to better utilization of power semiconductor devices.

22.3.4 Full-Bridge DC-DC Converter

Figure 22-13*a* shows the full-bridge dc-dc converter which is derived from the buck converter. The bridge circuit formed by switches S_1, S_2, S_3, and S_4 converts the input dc voltage to a high-frequency ac ($\geqslant$ 100 kHz), which is applied to the primary of transformer T_1. The high frequency results in a small size for the transformer. After isolation, the high-frequency ac at the secondary of the transformer is rectified by the center-tapped diode bridge rectifier formed by D_1 and D_2, and subsequently filtered by the *L* and *C* as in a buck converter. The topology is very popular for power levels greater than 500 W, when isolation is required.

Steady-state operating waveforms for the converter are shown in Fig. 22-13*b*. With switches S_3 and S_4 off, S_1 and S_2 are turned on simultaneously for $T_{on} = DT_s/2$, thereby applying a positive voltage across the transformer primary $T_{1,prim}$ and secondary $T_{1,sec1}$. During this time, diode D_1 conducts and a positive voltage appears across the inductor. With all the switches off, the transformer primary and secondary voltages are zero and the inductor current splits equally between diodes D_1 and D_2. In the second half of the switching cycle S_1 and S_2 are off, while S_3 and S_4 are simultaneously turned on $DT_s/2$. The rectified voltage waveform is similar to that in the buck converter and is at double the switching frequency of the each switch. The magnetizing flux in the transformer is bidirectional (Fig. 22-13*b*), resulting in better utilization of the core as discussed in *Transformers Design*. The conversion ratio is similar to the buck converter, but scaled by the secondary to primary transformer turns ratio.

22.3.5 Other Isolated DC-DC Converters

Several other isolated converters are based on the buck converter. Figure 22-14*a* shows the forward converter. The operation and conversion ratio is similar to the buck converter. However, the output voltage is scaled by the transformer turns ratio, an additional winding and diode (D_R) are needed to reduce the core flux to zero in each switching cycle, and an additional diode (D_2) is required at the output. The forward and flyback converters have unidirectional core flux and are limited to low-power applications. The push-pull converter (Fig. 22-14*b*), also derived from the buck converter, is better suited for higher power levels, limited by voltage rating of the switches required. Details of these converters can be found in Refs. [4, 5].

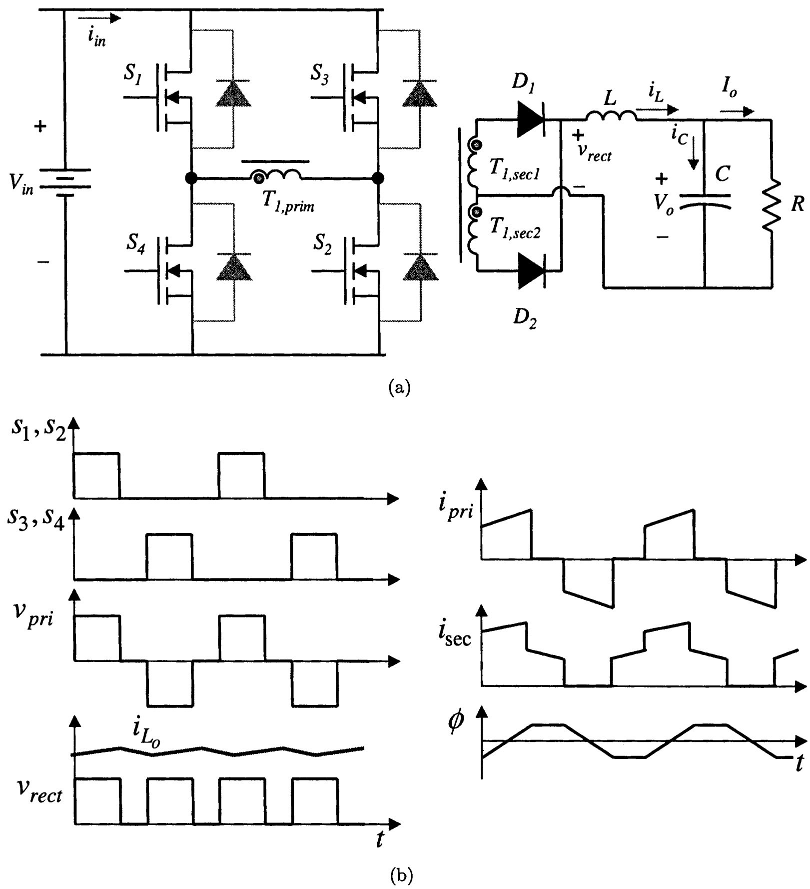


FIGURE 22-13 Full bridge dc-dc converter: (*a*) circuit, (*b*) steady-state waveforms.

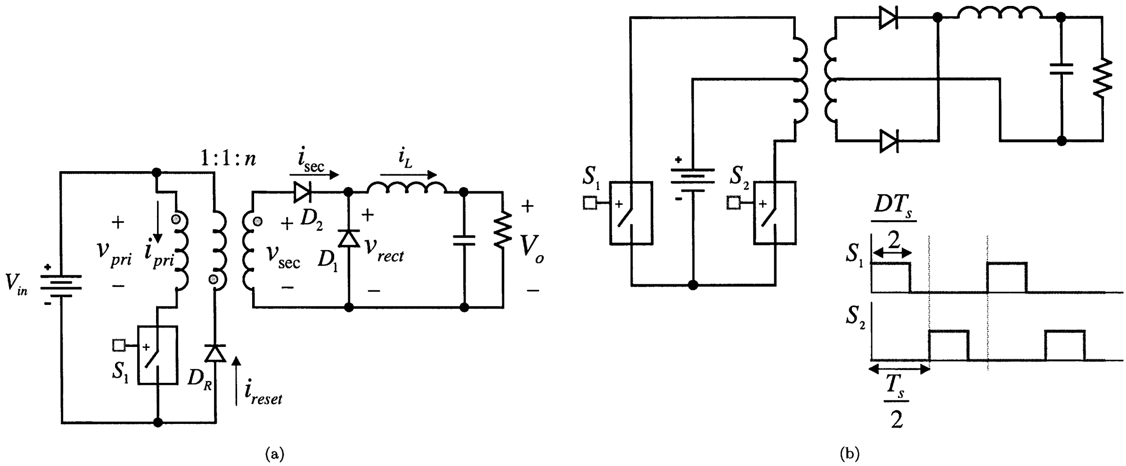


FIGURE 22-14 Other isolated converters: (*a*) Forward converter, (*b*) Push-Pull Converter.

22.3.6 Recent Developments and Future Trends

The size of filter components (inductor and capacitor) and isolation transformer reduce as the switching frequency is increased. Thus, a high switching frequency is desirable to minimize size and weight. However, parasitics and other nonidealities in dc-dc converters eventually limit the switching frequency and efficiency. For example, in flyback and forward converters, leakage inductance of the coupled inductor/transformer is a significant problem at high frequencies; in each switching cycle all the energy stored in the leakage inductance at switch turn-off has to be dissipated. Similarly, during turn-on of switches energy stored in parasitic output capacitance of the switch is dissipated inside the switch. These losses increase in proportion to the switching frequency. Thus, thermal or efficiency consideration eventually limit the maximum switching frequency. Besides the ones mentioned above other limiting factors are: switching times of diodes and MOSFETs, reverse recovery of diodes, capacitance of schottky diodes, and capacitance of transformers. To overcome these limitations, several circuits have been developed, which utilize the parasitic inductance and capacitances to advantage. Although the modifications in these sometimes add disadvantages, for specific applications, the advantages outweigh the disadvantages.

Soft-switching converters use resonance conditions between parasitic capacitance and inductance so that either the capacitance of switching devices is discharged before the device is actually turned on, or the current through the leakage inductance is reduced to zero prior to turning off. In some circuits, additional inductors and/or capacitors are added to produce resonance conditions. These converters, generically called *resonant converters*, are widely used in applications such as computer power supplies, electronic ballasts for fluorescent lamps, battery charging, and various portable applications. Details of these converters can be found in Refs. [4, 5].

Reduction in filtering requirements has also been achieved by using *interleaving*. An interleaved converter or multiphase converter has two or more converters called *phases*. These phases operate in parallel with their inputs and outputs being common. They are switched out of phase (180° for a 2-phase case, 120º for three phases, etc.) so that the ripple currents in the individual inductors are also out of phase. This results in lower effective current ripple both at the output and the input, and thus, smaller filter size for a given ripple specification. The lower values of the inductor also lead to faster dynamic response. Hybrid converters combining the benefits of soft switching with lower filter requirements have also been developed [13, 14].

To reduce size and minimize the number of discrete components, there is a significant effort in integrating all the semiconductors in one package. For example, on semiconductors [15] and power integration [16] have developed modules for use in off-line flyback converters; converters whose input is rectified line voltage are called off-line converters. The module contains a high-voltage power MOSFET and control circuit in one standard package. Similar modules are also available for low-power dc-dc converters (e.g., see Ref. [17]). Efforts are also being made to integrate all the magnetic components in dc-dc converters, using one single magnetic component, a concept called integrated magnetics.

22.4 FEEDBACK CONTROL OF POWER ELECTRONIC CONVERTERS

In the last section we saw that the steady-state output of a dc-dc converter, usually the output voltage, is controlled by the duty ratio. To account for changes in load current, input voltage, losses, and nonidealities in the converter, feedback based automatic control is required. Figure 22-15 shows a block diagram of output voltage control for a buck converter. The laplace domain control block diagram is also shown. The sensed output voltage is multiplied by a feedback gain $G_{FB}(s)$ before being compared with a reference value. The error is fed to an appropriate error compensator that generates a control voltage υ_c, which is converted to duty ratio d by the PWM block.

Toward designing a suitable controller, we will first describe a dynamic model of the power converter and then a simple loop-shaping control design method based on input to output bode plots.

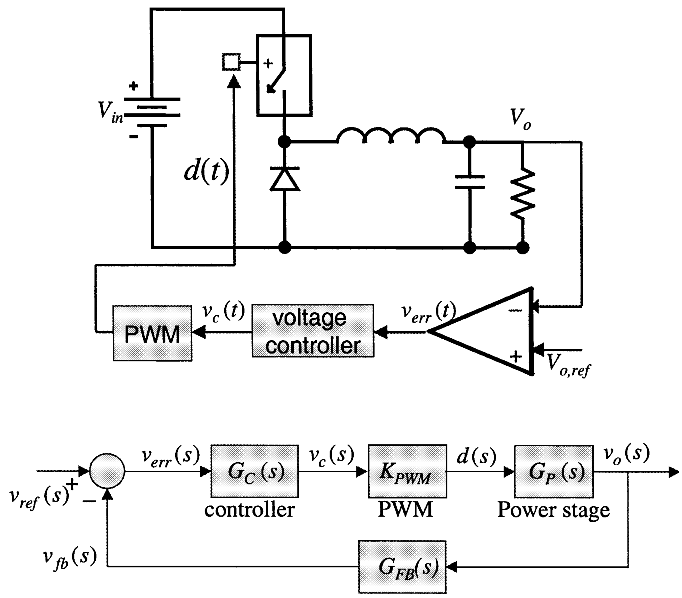


FIGURE 22-15 Block diagram of output voltage control for a buck converter.

It is possible to design more complex controllers in order to meet specific requirements, and the converter topology or operating method may also be modified to make the control design easier (e.g., DCM operation of flyback converters). To keep the explanation simple, it is assumed that the converter operates in CCM.

22.4.1 Dynamic Modeling

The power converter essentially consists of the PWM block and the power stage itself. The feedback gain G_{FB} is usually a constant. The PWM block shown in Fig. 22-10 converts the input control voltage $\bar{v}_c(t)$ to a duty ratio $d(t)$. From geometrical considerations

$$d(t)/\bar{v}_c(t) = 1/\hat{V}_{ramp} = K_R \tag{22-24}$$

where $\hat{V}_{ramp}$ is the peak value of the ramp $v_{ramp}(t)$. The power stage transfer function from $d(s)$ to $\bar{v}_o(s)$ can be derived using one of the methods stated below.

Dynamics of Averaged Quantities. As stated earlier, the bipositional switch approach and averaging are valid for analyzing low-frequency dynamics ($< f_s/5$) of the power converter. Unlike steady-state analysis, under dynamic conditions $\bar{v}_L \neq 0$ and $\bar{i}_C \neq 0$. Averaging the instantaneous state equations [Eq. (22-15)] over one switching cycle, dynamics of the averaged inductor current and capacitor voltage in a buck converter are

$$L\dot{\bar{i}}_L = d(t) \cdot V_{in} - \bar{v}_o \tag{22-25}$$

$$C\dot{\bar{v}}_o = \bar{i}_L - \bar{v}_o/R_L \tag{22-26}$$

Here the time varying duty cycle $d(t)$ is the control input and the averaged output voltage $\bar{v}_o$ is the output that has to be regulated. The situation for the buck converter is simple because the model described by Eqs. (22-25) and (22-26) is linear if V_{in} and R_L are assumed constant, for which case exact transfer functions describing large signal behavior can be derived. For boost and buck-boost converters, the averaged state equations involve terms with multiplications of $d(t)$ with a state variable. Thus, the model has to be linearized, and small signal dynamics obtained at different operating points are utilized for linear control design. It is of course possible to design large signal control based on the nonlinear model at the expense of mathematical complexity [18]. However, ease of design and simple cost-effective implementation has made linear design the preferred method in most power electronic converters in the low-to-medium power range.

Averaged Circuit Representation. Instead of writing averaged state equations explicitly as in Eqs. (22-25) and (22-26), an averaged circuit representation of the bipositional switch can be derived and substituted in different converter circuits Refs. [19–21]. As shown in Fig. 22-16*a* the bipositional switch can be considered as a two-port network with a voltage port (subscript$_{vp}$) at the input and a current port (subscript$_{cp}$) at the output. The average voltage and currents of the two ports are related as

$$\bar{v}_{cp}(t) = d(t) \cdot \bar{v}_{vp}(t) \tag{22-27}$$

$$\bar{i}_{vp}(t) = d(t) \cdot \bar{i}_{cp}(t) \tag{22-28}$$

The relations in the above equations correspond to those of an ideal transformer with turns ratio of $1 : d(t)$. Thus, for analysis purposes, the bipositional switch can be modeled as an ideal transformer whose turns ratio $d(t)$ can be controlled as shown in Fig. 22-16*b*. This representation is extremely useful in conjunction with circuit simulators which can perform operating point (dc bias) calculations, linearization, and ac analysis. Parasitic effects, like series resistances of inductors and capacitors,

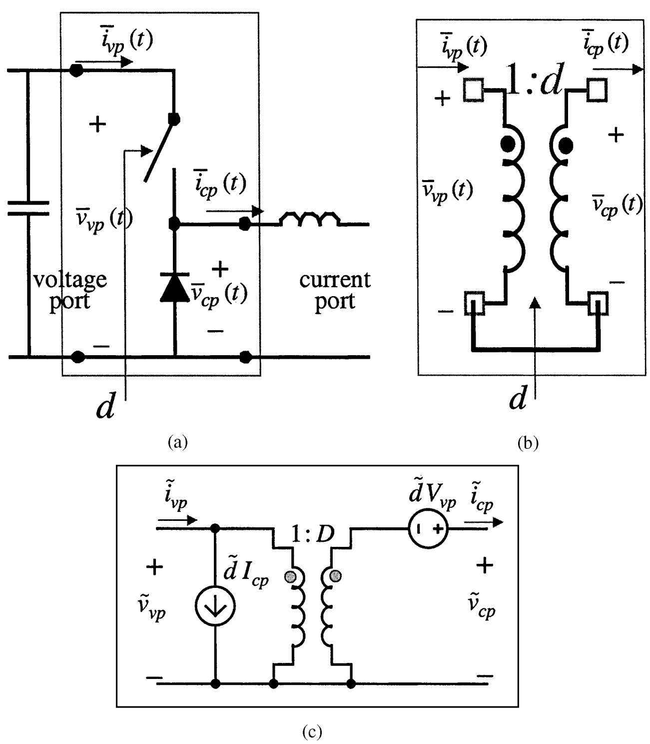


FIGURE 22-16 Bipositional switch: (*a*) two port network, (*b*) average representation, (*c*) small signal model.

can be easily incorporated in the averaged circuit model. Circuit simulators like SPICE [22], Saber, and Simplorer are commonly used for this purpose. The small-signal circuit representation of the averaged circuit obtained via linearization is shown in Fig. 22-16*c*. Quantities in upper case indicate operating point values, while the quantities with a "~" indicate small signal perturbations about the operating point. This representation can be utilized to derive small signal transfer functions using circuit analysis techniques.

22.4.2 Control Design

For a dc-dc converter, the main control objectives are: stability, zero steady-state error, specified transient response to step change in reference and in disturbance inputs (load and input voltage), and robustness to parametric changes. Transfer functions of different components—K_R, $G_{PS}(s)$, and $G_{FB}(s)$—are obtained as described above. The error compensator is then designed so that the open loop transfer function $G_{OL}(s)$ has a specified gain crossover frequency and phase margin. The gain cross-over frequency determines the response time of the controlled converter to changes in reference voltage and load current. Phase margin is usually in the range of 45° to 60° depending on the overshoot tolerable. Details on relation between gain crossover frequency, phase margin, and transient response can be found in any textbook on linear control (e.g., see Ref. 23).

An integrator (pole at origin) is added in $G_c(s)$ to obtain zero steady-state error. Zeroes are added at appropriate locations to obtain required phase margin. The dc gain of $G_c(s)$ is adjusted to achieve the required crossover frequency. Finally, to improve noise immunity, poles may be added for fast roll-off of the gain after the cross-over frequency. A systematic loop-shaping procedure along with implementation suited to dc-dc converters is described in Ref. [24].

Example: Voltage Control of a Buck Converter. A controller has to be designed to regulate the output voltage of a buck converter to a constant value. The specifications, parameters, and control requirements are listed in Table 22-2. Using the methods described above, the duty ratio to output voltage transfer function can be derived to be

$$\frac{\tilde{v}_o(s)}{\tilde{d}(s)} = \frac{V_{in}(1 + sCR_{ESR})}{1 + s[CR_{ESR} + L/R_L] + s^2LC[1 + R_{ESR}/R_L]} \tag{22-29}$$

The transfer function has a complex pole pair due the L-C filter, and a left half zero due to the ESR of the output capacitor. The compensator designed in accordance with the aforementioned considerations is

$$G_c(s) = \frac{K_c(1 + s/\omega_z)^2}{s(1 + s/\omega_{p1})(1 + s/\omega_{p2})} \tag{22-30}$$

where $K_c = 2011$, $\omega_z = 2\pi \times 2556$ rad/s, $\omega_{p1} = 2\pi \times 11.3e3$ rad/s, and $\omega_{p1} = 2\pi \times 80e3$ rad/s.

Representation of the controlled converter using ORCAD PSpice [25] is shown in Fig. 22-17*a*. The bipositional switch has been replaced by a two-port network that models a transformer with

TABLE 22-2 Control Design Example

Specifications		Parameters		Requirements	
Input voltage	20 [V] to 30 [V] dc	L	75 [μH]	Cross-over frequency	8 [kHz]
Output voltage	15 [V] dc	C	47 [μF]	Phase margin	> 60°
Maximum load current	5 [A]	R_{ESR}	0.3 [Ω]		
Switching frequency	200 [kHz]	K_R	1		

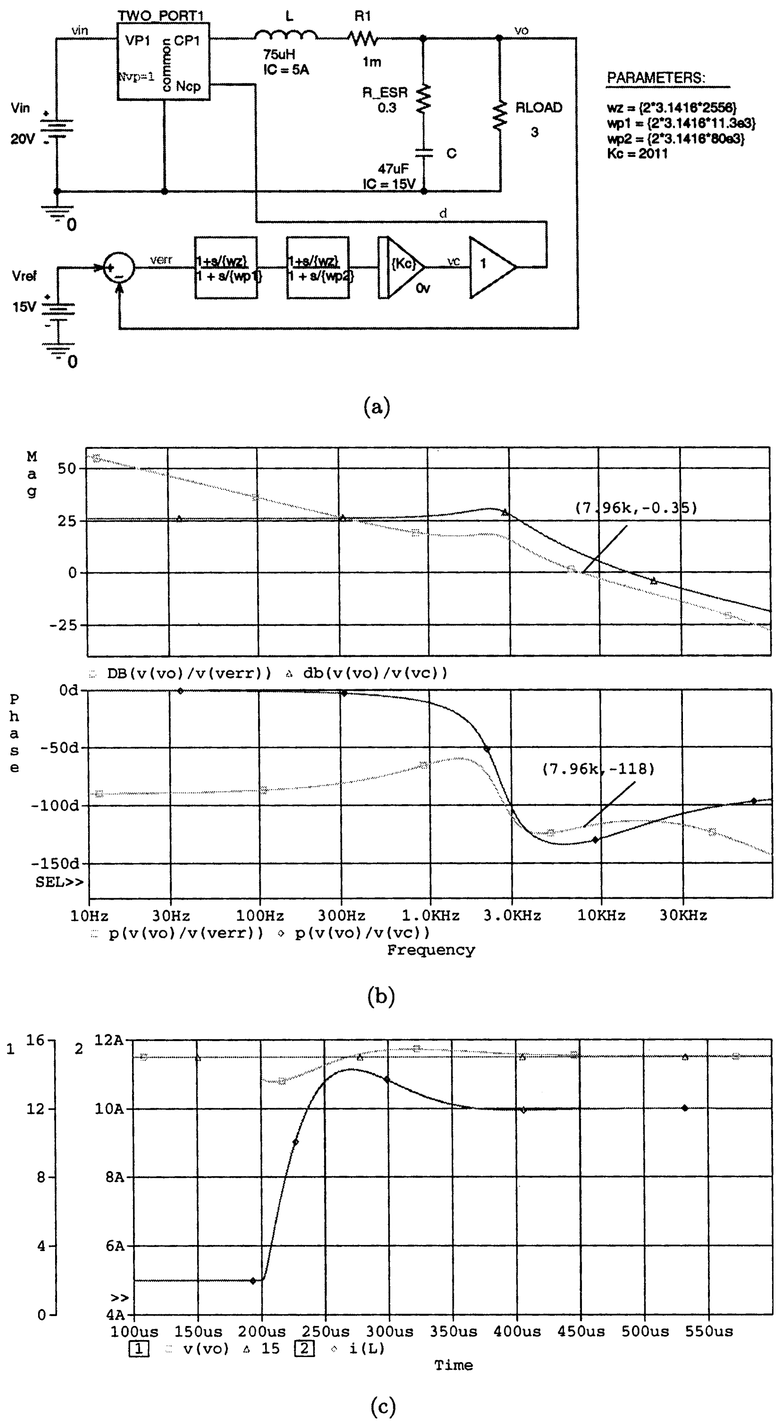


FIGURE 22-17 Controlled buck converter: (*a*) averaged representation using PSpice, (*b*) open loop bode plots, (*c*) transient response.

controllable turns ratio. The compensated and uncompensated transfer functions obtained using ac analysis are shown in Fig. 22-17*b*. The gain crossover frequency and the corresponding phase of the compensated transfer function are indicated. Figure 22-17*c* shows dynamic response of the controlled converter when a step change in load is applied at 0.2 ms.

22.4.3 Current Mode Control

In most converters, the inductor current is an *internal* state of the power converter. Changes in the input voltage and duty ratio are first reflected in the inductor current, and subsequently in the output voltage. Thus, controlling the inductor current can lead to better performance. Figure 22-18 shows a cascaded control structure where the internal current controller is about an order of magnitude faster than the outer voltage loop. The average value of the inductor current is controlled to a reference that is generated by the error compensator for the voltage-control loop. The current controller is designed using the transfer function from the duty ratio to the inductor current. For voltage controller, the current control loop is assumed to be ideal, that is, $i_L = i_{L,ref}$; this is justified since the current controller is much faster than the voltage controller. The voltage compensator is then designed, using the inductor current to the output voltage transfer function. In a buck or buck-derived topology, the average inductor current is equal to the load current. Thus, fast control over the inductor current effectively mitigates steady state and transient variations in the input voltage without affecting the output voltage. This current control method is called *average current control*.

Another popular method is *peak current mode control*. In this method, the peak value of the inductor current is controlled to the reference value (generated by the voltage-control loop) in each switching cycle. Peak current mode control has the additional advantage of balancing the positive and negative flux excursions in transformer isolated topologies like full bridge and push pull. However, peak current mode control requires extra precautions to avoid subharmonic and chaotic operation [4, 5, 26].

22.4.4 Other Control Techniques

The basic modeling method described above is applicable to other types of converters (like dc-ac) as long as low-frequency behavior is being studied. Dynamics of the filter elements may of course be different. In dc-ac converters, the control objective is usually to track a moving reference (e.g., sinusoidal control voltage or current). Using a stationary to rotating frame transformation, commonly called the *abc to dq transformation*, the tracking problem can often be reduced to a regulation problem. If current control is implemented in the stationary frame, where the control objective is to track a sinusoidally varying reference, then either average current control or *hysteretic current control* is used. In hysteretic current control, the current is maintained in a band about the reference

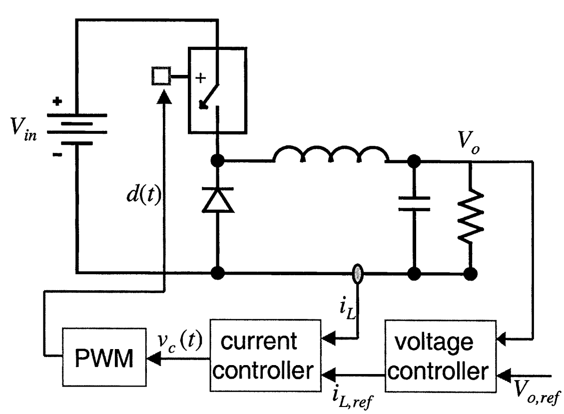


FIGURE 22-18 Average current control of a buck converter.

value. If the current error is below the lower limit of the band, a positive voltage is applied across the inductor (switch on in a buck converter) to increase the current. To reduce the current a negative or zero voltage is applied across the inductor. With hysteretic control, the rise and fall times are only limited by the power components of the converter. However, it has the disadvantage of variable switching frequency, whose instantaneous value depends on a combination of several factors. Several other techniques have recently been proposed for control of power electronic converters: sigma-delta, sliding mode, dead beat, etc. Details of these control techniques can be found in Ref. [27]. In digital implementations, predictive current control is commonly used to reduce the effect of sensing and computational delays. So far, digital control is only used in high power converters, where the overall cost justifies the cost of digital and interface components. However, there is significant effort in extending the benefits of digital control to lower power converters.

22.5 DC-AC CONVERSION: INVERSION

DC-ac converters constitute a significant portion of power electronic converters. These converters, also called *inverters*, are used in applications such as electric motor drives, uninterruptible power supplies (UPS), and utility applications such as grid connection of renewable energy sources. Inverters for single phase ac and 3-phase 3-wire ac systems (without a neutral connection) are described in this section.

22.5.1 Single Phase AC Synthesis

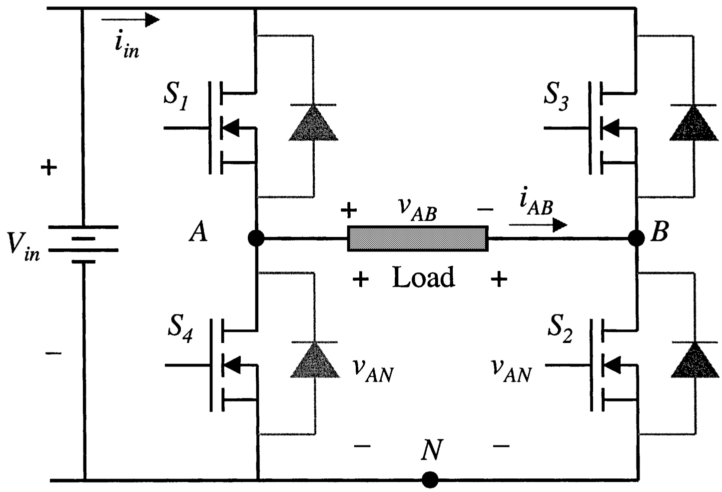
In an ac system both the voltage and the current should be able to reverse in polarity. Further, the voltage and current polarities may or may not be the same at a given time. Thus, a dc-ac converter implementation should be able to output a voltage independent of current polarity. In the full-bridge dc-dc converter shown in Fig. 22-19*a*, the primary circuit consisting of four controlled switches, also called *H-bridge*, has two bipostional switch implementations. Each bipositional switch has bidirectional current capability but only positive output voltage ($v_{AN} >, v_{BN} > 0$). However, based on the duty cycles, the difference of the outputs, $V_{AB} = v_{AN} - v_{BN}$, can reverse in polarity. Thus, the H-bridge is used for synthesizing single phase ac voltage from a dc voltage.

Quasi-square Wave Inverter. The simplest form of dc-ac conversion, albeit with poor quality, is synthesis of quasi-square wave ac instead of a pure sine wave. Diagonally opposite switches in the H-bridge are turned on simultaneously. The pulse width of each pair is controlled to adjust the magnitude of the fundamental component, while the switching frequency is equal to the required output frequency. The synthesized voltage waveform is shown in Fig. 22-19*b*. The peak value of fundamental and harmonic components are

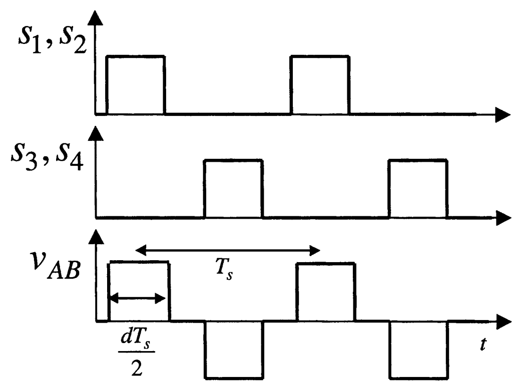
$$V_{AB,n} = \frac{4V_{in}}{n\pi}\sin(n\pi d/2) \qquad n \text{ odd} \tag{22-31}$$

where d is the duty ratio and n is the harmonic number. This converter is widely used for low cost low power UPS applications where the voltage waveform quality is not important. Incandescent lighting, universal input motors, and loads with a diode bridge or power factor corrected front end (discussed in Sec. 22.8.1) are not affected by the voltage waveform quality. The load current, i_{AB}, has harmonics based on the load characteristics. Sometimes an LC filter is added at the output to reduce the voltage (and therefore the current) harmonics.

Single-Phase Sinusoidal Voltage Synthesis. For applications requiring low voltage and current distortion, high-frequency PWM is utilized to generate a sinusoidally varying average voltage. The power converter used is the H-bridge shown in Fig. 22-19*a*. The duty ratio for each bipositional switch, also called one leg of the inverter, is varied sinusoidally. The switching signals are generated



(a)



(b)

FIGURE 22-19 Single-phase inverter: (*a*) circuit, (*b*) quasi-square wave synthesis.

by comparison of a sinusoidally varying control voltage with a *triangle* wave as shown in Fig. 22-20. Equations relating the control voltages, duty ratios, and the averaged output voltages are as follows:

$$v_c = \hat{V}_c \cdot \sin(\omega_m t) \tag{22-32}$$

$$v_{cA}(t) = v_c = \hat{V}_c \cdot \sin(\omega_m t) \tag{22-33}$$

$$v_{cB}(t) = -v_c = -\hat{V}_c \cdot \sin(\omega_m t) \tag{22-34}$$

$$d_A(t) = \frac{1}{2}\left(1 + \frac{v_{cA}}{\hat{V}_{tri}}\right) = \frac{1}{2}(1 + m \cdot \sin(\omega_m t)) \tag{22-35}$$

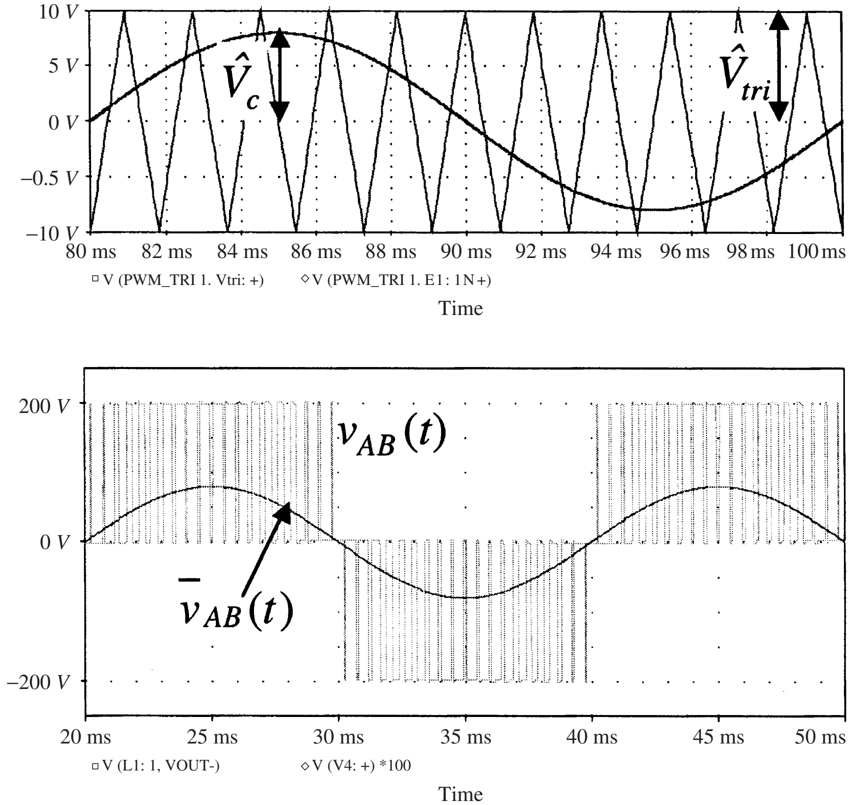


FIGURE 22-20 Single-phase sinusoidal ac synthesis waveforms.

$$d_B(t) = \frac{1}{2}\left(1 + \frac{v_{cB}}{\hat{V}_{tri}}\right) = \frac{1}{2}(1 - m \cdot \sin(\omega_m t)) \tag{22-36}$$

$$\bar{v}_{AN}(t) = d_A(t) \cdot V_{in} = \frac{1}{2}(1 + m \cdot \sin(\omega_m t)) \cdot V_{in} \tag{22-37}$$

$$\bar{v}_{BN}(t) = d_B(t) \cdot V_{in} = \frac{1}{2}(1 - m \cdot \sin(\omega_m t)) \cdot V_{in} \tag{22-38}$$

$$\bar{v}_{AB}(t) = (d_A(t) - d_B(t)) \cdot V_{in} = m \cdot V_{in} \cdot \sin(\omega_m t) = \underbrace{(V_{in}/\hat{V}_{tri})}_{k_{PWM}} \cdot v_c(t) \tag{22-39}$$

Here $\hat{V}_c$ and V_{tri} are peak values of control voltage and the triangle wave, respectively, $m = \hat{V}_c/\hat{V}_{tri} \in [0, 1]$ is the modulation index, $\omega_m = 2\pi f_m$ is the angular frequency of the sinusoid to be synthesized, while $d_A(t)$ and $d_B(t)$ are duty ratios of switches S1 and S3, respectively. In Eq. (22-39) k_{PWM} may be regarded as the gain of the power converter that amplifies the control signal $v_c(t)$ to the average output voltage $\bar{v}_{AB}(t)$. The maximum peak value of the output voltage, obtained for $m = 1$, is V_{in}. This is significantly lower than that obtainable with the quasi square wave inverter ($4V_{in}/\pi$). However, harmonics in the output voltage are significantly reduced and are at much higher frequencies: $k \cdot f_s \pm l \cdot f_m$, where k and l are integers such that $k + l$ is odd [4]. The switching frequency $f_s \geqslant 20$ kHz is significantly higher than the output frequency f_m, which usually has a maximum value of about 50/60 Hz. If the load is inductive, the current harmonics are reduced further, and the current is almost sinusoidal.

Equation (22-37) can be rewritten as

$$\overline{v}_{AN}(t) = d_A(t) \cdot V_{in} = \frac{V_{in}}{2} + \frac{k_{PWM}}{2} \cdot v_c(t) \tag{22-40}$$

This clearly shows that on an average basis the "neutral point" for the output of one inverter leg is $V_{in}/2$ above "N," that is, at the mid-point of the input dc bus. Thus, using the same H-bridge a split-phase ac (two ac voltages 180° out of phase with a common return) can be generated if the center point of the dc bus is available as the neutral connection for the output.

22.5.2 Three-Phase AC Synthesis

The last observation in the previous section leads us to 3-phase inverters without a neutral connection. The circuit consists of three legs, one for each output with a common dc link as shown in Fig. 22-21*a*.

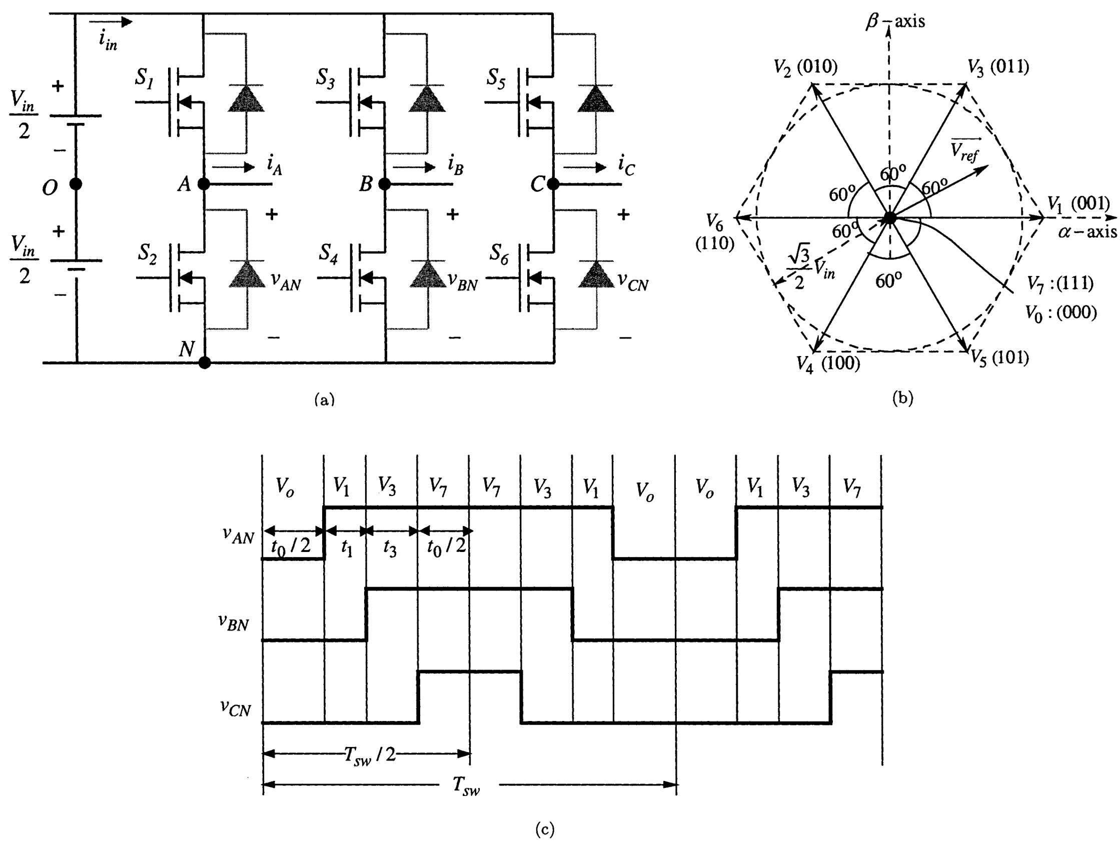


FIGURE 22-21 3-Phase ac synthesis: (*a*) converter, (*b*) output voltage vectors, (*c*) instantaneous waveforms.

Using sine triangle PWM with control voltages offset by 120° (instead of 180° as in the single phase case) we obtain

$$v_{cA}(t) = \hat{V}_c \cdot \sin(\omega_m t) \tag{22-41}$$

$$v_{cB}(t) = \hat{V}_c \cdot \sin(\omega_m t - 2\pi/3) \tag{22-42}$$

$$v_{cC}(t) = \hat{V}_c \cdot \sin(\omega_m t + 2\pi/3) \tag{22-43}$$

$$\overline{v}_{AN}(t) = \frac{V_{in}}{2} + \frac{k_{PWM}}{2} \cdot v_{cA}(t) \tag{22-44}$$

$$\overline{v}_{BN}(t) = \frac{V_{in}}{2} + \frac{k_{PWM}}{2} \cdot v_{cB}(t) \tag{22-45}$$

$$\overline{v}_{CN}(t) = \frac{V_{in}}{2} + \frac{k_{PWM}}{2} \cdot v_{cC}(t) \tag{22-46}$$

The zero-sequence component of the output voltages, $v_z = (v_{AN} + v_{BN} + v_{CN})/3 = V_{in}/2$, does not appear in the line-to-line voltages, and since there is no neutral connection to the inverter zero-sequence currents do not flow.

The maximum peak value of the output line-to-line voltages is $\hat{V}_{LL} = (\sqrt{3}/2)V_{in}$. Using square wave inversion, similar to that for the single-phase case, we can obtain higher magnitude for the fundamental component of the output voltages at the cost of adding harmonics. However, if, instead of all the harmonics, only the fundamental and those harmonics of the square wave that contribute zero-sequence component (triplen harmonics) are retained, the output voltage amplitude increases without adding harmonics to the line-to-line voltages and the line currents. Usually, addition of the third harmonic component is sufficient Refs. [28, 29]. As described in Refs. [30, 31], the most effective method is to add the following zero-sequence component to the control voltages for each phase

$$v_{cz}(t) = \frac{1}{2}[\max(v_{cA}(t), v_{cB}(t), v_{cC}(t)) + \min(v_{cA}(t), v_{cB}(t), v_{cC}(t))] \tag{22-47}$$

In terms of output voltage generation, this is equivalent to space vector modultation (SVM).

22.5.3 Space Vector Modulation

This method has become extremely popular for 3-phase inverters in the low-to-medium power range. A very brief description will be presented here and details can be found in Refs. [27, 28, 30].

For 3-phase systems with no zero-sequence component, that is, $v_z = (v_{AN} + v_{BN} + v_{CN})/3 = 0$, the 3-phase quantities are linearly dependent and can be transformed to a 2-phase orthogonal system commonly called the $\alpha\beta$ system. Quantities in the $\alpha\beta$ system can be represented by complex numbers and as two-dimensional vectors in a plane, called *space vectors*. The transformation from the *abc* to $\alpha\beta$ quantities is given by

$$\vec{v}_{\alpha\beta}(t) = v_\alpha(t) + j \cdot v_\beta(t) = e^{j0} \cdot v_a(t) + e^{j2\pi/3} \cdot v_b(t) + e^{j4\pi/3} \cdot v_c(t) \tag{22-48}$$

With negative sequence components absent, α and β components of steady-state sinusoidal *abc* quantities are also sinusoids with constant amplitude and a 90° phase difference between them. Under transient conditions they are arbitrary time-varying quantities. Thus, for balanced sinusoidal conditions, the space vector $\vec{v}_{\alpha\beta}(t)$ rotates in counter clockwise direction with angular frequency equal to frequency of the *abc* voltages, and describes a circle of radius $(3/2)\hat{V}_{ph}$, $\hat{V}_{ph}$ being the peak of the phase voltage.

The instantaneous output voltages of the 3-phase inverter shown in Fig. 22-21*a* can assume eight different combinations based on which of the six MOSFETs are on. The space vectors for these eight combinations are shown in Fig. 22-21*b*. For example, vector V_4 denoted by (100) corresponds to switch states $\upsilon_{AN} = V_{in}$, $\upsilon_{BN} = 0$, and $\upsilon_{CN} = 0$. The vectors $V_0(000)$ and $V_7(111)$ have zero magnitude and are called *zero vectors*.

Synthesis utilizing the idea of space vectors is done by dividing one switching time period into several time intervals, for each of which a particular voltage vector is the output by the inverter. These time intervals and the vectors applied are chosen so that the average over one switching time period is equal to the desired output voltage vector. For the reference voltage vector $\vec{\upsilon}_{ref}$, shown in Fig. 22-21*b*, the nonzero vectors adjacent to it (V_1 and V_3), and the zero vectors (V_0 and V_7) are utilized as shown in Fig. 22-21*c*. Relative values of time intervals t_1 and t_3 determine the direction, while ratio of t_0 to the switching time period determines the magnitude of the output vector synthesized.

The maximum obtainable average vector lies along the hexagon connecting the six nonzero vectors. As stated earlier, balanced 3-phase sinusoidal quantities describe a circle in the $\alpha\beta$ plane. Thus, to synthesize distortion-free and balanced 3-phase sinusoidal voltages the circle must be contained within the hexagon, that is, with a maximum radius of $\sqrt{3}/2 \cdot V_{in}$. This gives the maximum peak value of line-to-line voltage obtained with SVM as $\hat{V}_{LL} = V_{in}$. This is significantly higher than that obtained using sine triangle PWM: $\sqrt{3}/2 \cdot V_{in}$. Further, the sequence and choice of vectors applied can be optimized to minimize number of switchings and ripple in the resulting currents [32]. There are several variations of SVM, each suited to a different application. Space vector modulation can be easily implemented digitally using microcontroller and digital signal processor (DSPs), and is extremely advantageous in control of 3-phase ac machines strategies using *vector control* and *direct torque control* (DTC) [33–37].

22.5.4 Multilevel Converters

The converter topologies described so far are based on a 2-level converter leg (bipositional switch), where the output voltage of each leg (υ_{AN}) can be either zero or V_{in}. The converters are therefore called 2-level converters. In 2-level converters, all the switches have to block the full dc bus voltage (V_{in}). For high-power applications insulated gate bipolar transistors (IGBTs) and gate turn-offs (GTOs) are used as the semiconductor switches. These have higher voltage and current ratings, and lower on-state voltage drop compared to power MOSFETs, but cannot switch as fast. In some applications like some motor drives and utility applications, even the voltage ratings of available IGBTs and GTOs is not sufficiently high. Simple series connection, to achieve a higher blocking voltage, has problems of steady state and dynamic voltage sharing. Moreover, due to the low switching frequency of high-power switches, the output voltage and current quality deteriorates. These issues are addressed by multilevel converters. In a multilevel converter [38, 39], the output of each phase leg can attain more than two levels leading to improved quality of the output voltage and current. The circuit comprising each leg and its proper operation ensure that voltage blocked by the switches reduces as the number of levels is increased. In addition, multilevel converters are modular to some extent, thereby making it easy to scale voltage ratings by increasing the number of "cells".

Multilevel PWM. For 2-level PWM, comparison of the control voltage with a triangle wave generates the switching signal for the top switch, while the bottom switch is controlled in complement to the top switch. Each of these two states corresponds to the two levels of the output voltage. For multilevel converters, there are more than two effective switch states, each of which corresponds to an output voltage level. For example, in a 3-level converter there are three effective states $q(t) = 0$, 1, 2, corresponding to output voltage levels $\upsilon_{AN}(t) = 0$, $V_{in}/2$, V_{in}. The control voltage $\upsilon_c(t)$ is compared with two triangle waves to obtain two switching signals $q_1(t)$ and $q_2(t)$, and the effective switching signal can be obtained as $q(t) = q_1(t) + q_2(t)$ as shown in Fig. 22-22. The output voltage is then given by $\upsilon_{AN} = q(t) \cdot (V_{in}/2)$. Switching signals for the individual switches are derived using $q(t)$ and the circuit topology. For the waveforms in Fig. 22-22, $f_s = 60Hz$ and $V_{in} = 2\ kV$. Since the

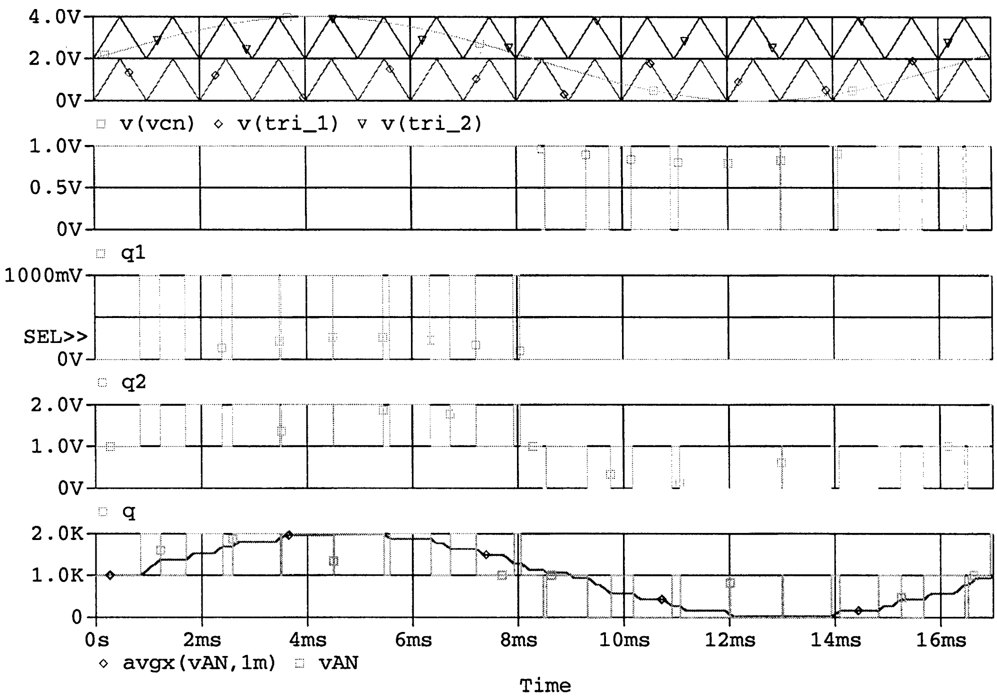


FIGURE 22-22 Multilevel triangle comparison.

v_{AN} waveform is closer to desired sinusoid in the 3-level case, the output voltage has lower total harmonic distortion (THD) even if the switching frequency is low. For 3-phase converters, space vector–based PWM can be used for generating the switching signals (e.g., see Ref. 40), the advantage in the multilevel case compared to the 2-level case being the significantly higher number of output voltage vectors.

Multilevel Converter Topologies. There are three basic multilevel converter topologies—diode clamped, flying capacitor, and cascaded full-bridge converters.

Diode Clamped Converter. Figure 22-23*a* shows 1-phase leg of a 3-level diode clamped converter [41]. The input dc bus is split by means of capacitors. Pairs of switches are turned on to obtain three different voltage levels for the output voltage $v_{AN} = 0$, $V_{in}/2$, V_{in}, as shown in Fig. 22-23*c*. It is evident that this circuit acts like a *tripositional* switch connecting the output to one of three positions of the input dc bus. The minimum voltage at point b_1, and the maximum voltage at point b_2, is clamped to $V_{in}/2$ by the blocking diodes D_{b1} and D_{b2}, respectively. Thus, all the switches have to block $V_{in}/2$ during their off state. This topology can be extended to more number of levels. However, it is eventually limited by the voltage rating of blocking diodes, which have to block increasing voltages as the number of levels is increased. One-phase leg of a 5-level version is shown in Fig. 22-23*b*.

Flying Capacitor Converter. Figure 22-24 shows the topology of a 3-level flying capacitor converter. The basic idea here is that the capacitor *C* is charged to half the input dc voltage by appropriate control of the switches. The capacitor can then be inserted in series with the output voltage—either adding or subtracting $V_{in}/2$, and thereby giving 3-output voltage levels.

Cascaded Full Bridge Converters. In this scheme [42], single-phase H-bridges shown in Fig. 22-19*a* are connected in series at the output to form one single phase circuit. Three separate

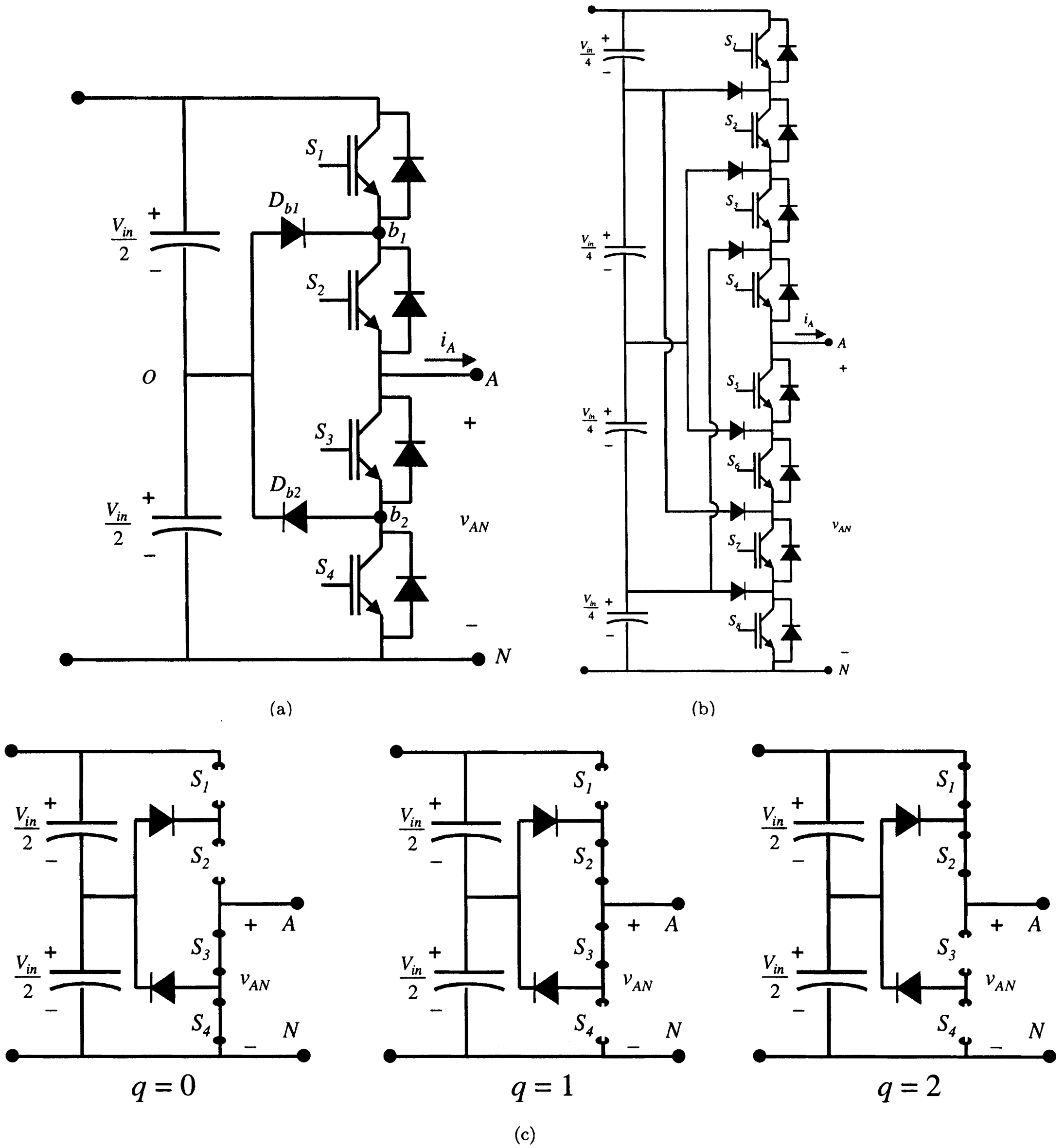


FIGURE 22-23 Diode clamped converters: (*a*) one phase of a 3-level converter, (*b*) one phase of 5-level converter, (*c*) switching states in a 3-level converter.

circuits are required for a 3-phase implementation. Since all the H-bridges are same, the circuit is modular and can be scaled by adding more H-bridges. However, dc sources at the input of all H-bridges have to be isolated from each other. It is also possible to combine different types of H-bridges—IGBT-based fast switching type and GTO-based slower switching type—or have different dc bus voltage magnitudes in different bridges to optimize losses or increase effective number of levels.

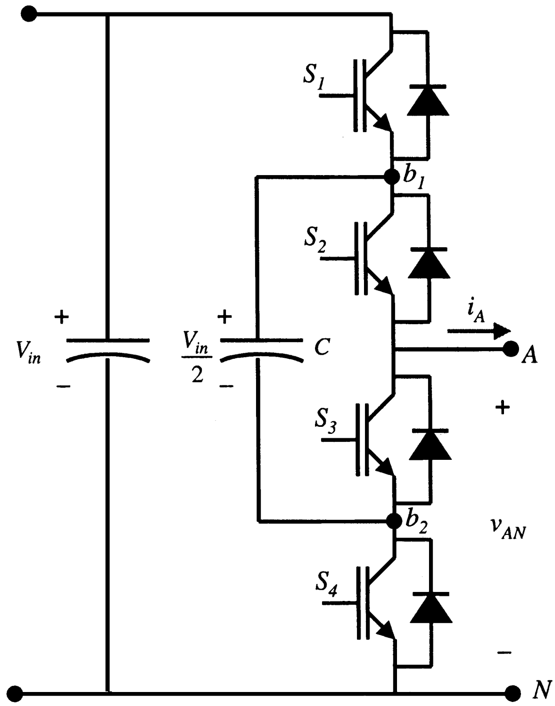


FIGURE 22-24 3-Level flying capacitor converter.

22.6 AC-DC CONVERSION: RECTIFICATION

AC-dc converters, or rectifiers, are used at the input of almost all line connected electronic equipment. Electronic devices that are powered directly from line and do not have regulation requirements use single- and 3-phase diode bridge rectifiers for converting line frequency ac to an uncontrolled dc voltage. For control over the output dc voltage thyristor-based rectifiers are used. Power factor corrected front end converters, discussed in Sec. 22.8, provide output voltage regulation as well as near unity power factor.

22.6.1 Single-Phase Diode Bridge Rectifier

Figure 22-25*a* shows the circuit of a single-phase diode bridge rectifier with a purely capacitive output filter. Due to its simplicity and low cost this circuit is preferred for low-power applications such as input stages of ac-dc adapters and computer power supplies. Diodes conduct in pairs to transfer energy from the input to the output when the input line voltage exceeds the output dc voltage in magnitude. Diodes D_1 and D_4 conduct when $v_s > v_o$, while D_2 and D_4 conduct when $-v_s > v_o$. The capacitor C_d gets charged by high current pulses during these small intervals near the peak of v_s, and discharges with the almost constant load current during the rest of the line cycle, as shown in Fig. 22-25*b*. The output dc voltage is approximately equal to the peak of the line voltage minus the forward voltage drop of two diodes. The capacitor value is chosen on the basis of the maximum load current and allowable output voltage ripple. The line current has significant harmonic content as shown in Fig. 22-25*c*. Source inductance of the line, common for regular utility supply, leads to lower peak input current, larger conduction times for the diodes, and reduced magnitude of the output voltage.

To quantify the line current distortion the following definitions are commonly used.

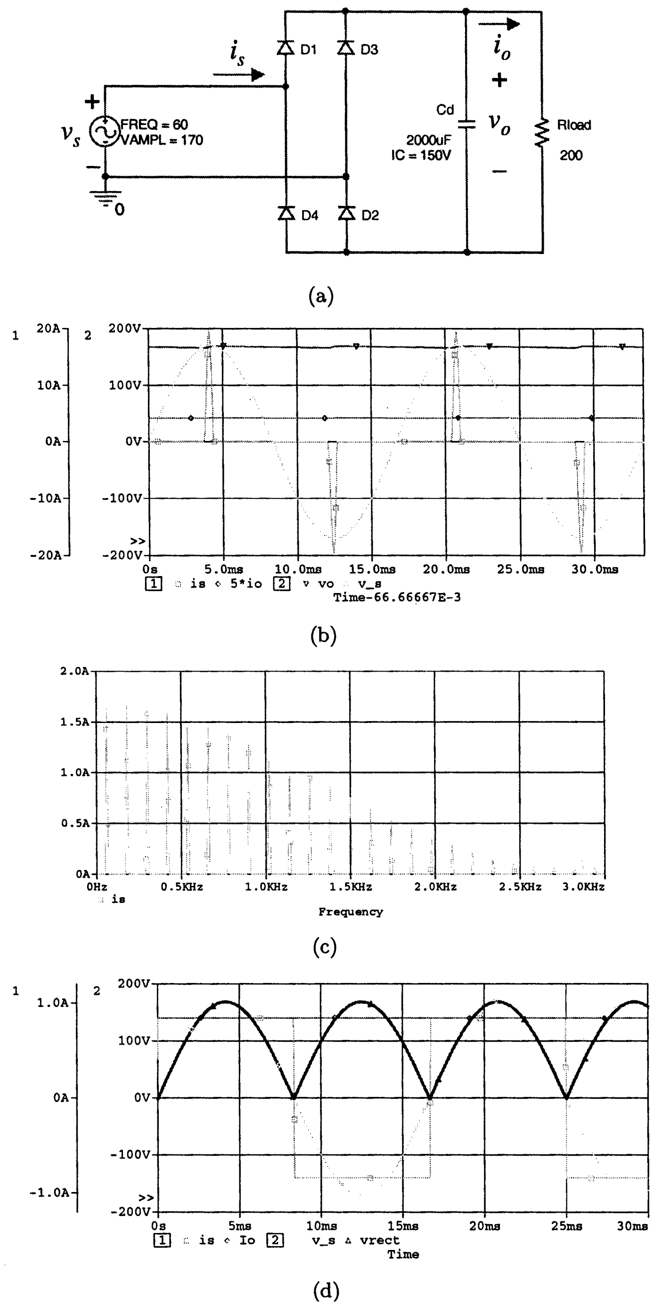


FIGURE 22-25 Single-phase diode bridge rectifier (*a*) circuit, (*b*) waveforms, (*c*) line current harmonics, (*d*) waveforms with inductive filter.

Total Harmonic Distortion. THD is the ratio of rms values of the distortion component to the fundamental component, expressed as a percentage

$$\mathrm{THD} = \frac{I_{dist}}{I_1} \times 100 = \frac{\sqrt{(I^2 - I_1^2)}}{I_1} \times 100 \tag{22-49}$$

Real Power. This is the actual value of power consumed computed as an average over one line cycle

$$P_{real} = \frac{2\pi}{\omega}\int_0^{2\pi/\omega} v_s(t)i_s(t)\,dt = VI_1\cos(\phi_1) \tag{22-50}$$

Apparent Power. It is the product of the rms values of the input voltage and current

$$P_{app} = V \cdot I \tag{22-51}$$

Power factor. Power factor (PF) is defined as the ratio of real power to apparent power.

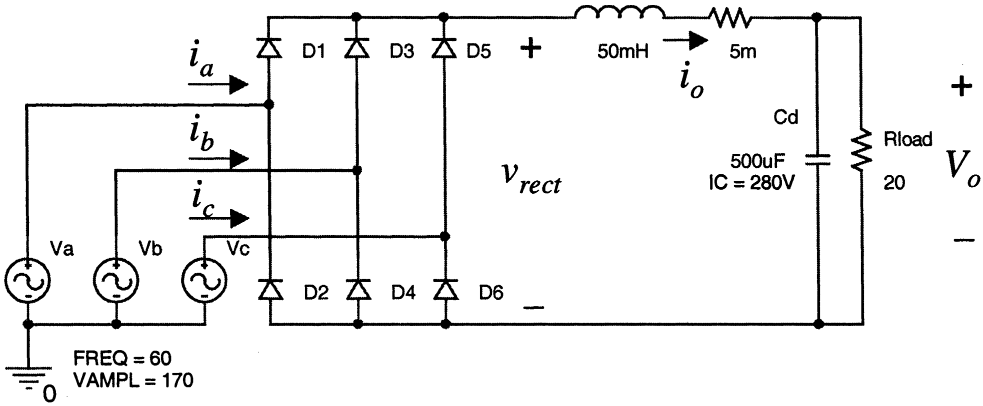
$$PF = \frac{P_{real}}{P_{app}} = \frac{VI_1\cos(\phi_1)}{VI} = \frac{I_1}{I} \cdot \cos(\phi_1) \tag{22-52}$$

where V, I, and I_1 denote the rms value of the voltage, current, and fundamental component of the current, respectively, ϕ_1 is the phase angle of the fundamental component of the current with respect to the input voltage (assumed sinusoidal), and I_{dist} is the rms value of the distortion component of the input current. The term $\cos(\phi_1)$ is called the *displacement power factor*, while the term I_1/I is called the *distortion power factor*.

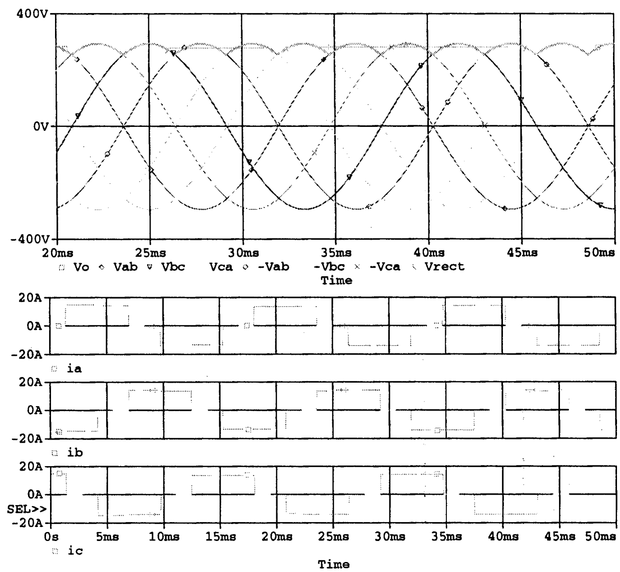
For the circuit values of Fig. 22-25*a*, the load current is 0.84 A, the peak line current $\hat{I}_s = 19.4$ A, rms line current $I_s = 3.8$ A, rms of the fundamental component $I_{s1} = 1.18$ A, $THD = 280\%$, and the PF = (1.18/3·8) · cos(2.0°) = 0.31. The quality of the input current can be improved significantly if an inductive filter is used at the output of the rectifier. With a high enough inductance, the output current can be maintained nearly constant. This leads to a square wave shape for the input current as shown in Fig. 22-25*d*, which has a THD of 48% and a PF of 0.9. With the inductive filter, the output voltage has an average value equal to the average value of a rectified sinusoid, that is, $2\hat{V}_s/\pi$, where V_s is the peak value of the input phase voltage. Inductive output filter is preferable for medium power applications so that the input current has lower harmonic content.

22.6.2 Three-Phase Diode Bridge Rectifier

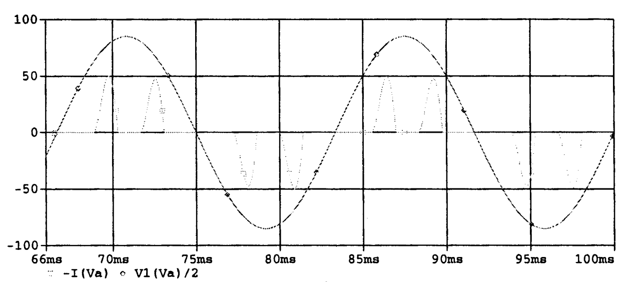
Figure 22-26*a* shows a 3-phase diode bridge rectifier with an inductive output filter. The operation is similar to the single-phase case. Diodes conduct in pairs—one from the upper three and one from the lower three. Cathodes of diodes D_1, D_3, and D_5 are connected together, so the diode with the highest voltage at its anode conducts. The converse holds for diodes D_2, D_4, and D_6. The rectified voltage follows the envelope of the line voltages and their negatives: $v_{rect} = \max(|v_{ab}|, |v_{bc}|, |v_{ca}|)$. This rectifier is also called the 6-pulse rectifier because the voltage at the output of the diode bridge, v_{rect}, has six pulses in every line cycle. The average output voltage across the load is $V_o = \bar{v}_{rect} = (3/\pi)\hat{V}_{LL}$, $\hat{V}_{LL}$ being the peak value of the line-to-line voltage. The input line currents can be derived considering which diodes are conducting at a given time. They have quasi-square waveshapes as shown in Fig. 22-26*b*. The harmonic distortion is lower than in the single-phase case with inductive filter: THD = 31% and PF = 0.955. If the output filter is purely capacitive, the output voltage is equal to $\hat{V}_{LL}$, while the input currents are significantly distorted (Fig. 22-26*c*) and have harmonics at $(6m\pm1)f$, where f is the line frequency and m is an integer. As in the single-phase rectifier with capacitive output filter, THD of the input current depends significantly on the source impedance.



(a)



(b)



(c)

FIGURE 22-26 3-Phase diode bridge rectifier (*a*) circuit, (*b*) waveforms with inductive filter, (*c*) line current waveform with capacitive filter.

The quality of input current and PF generally improve when going from single-phase to three phase, and can be further improved with higher number of phases if voltages with appropriate phase difference are generated from the utility supplied three phases. The output filter requirements also reduce as the number of phases is increased. With six voltage sources phase shifted by 30°, 12 diodes can be utilized to generate a 12-pulse rectifier. Isolated voltage sources phase shifted by 30° can be obtained using a wye-delta connected 3-phase transformer. Other phase shifts are generated by vectorial combination of appropriately scaled and isolated voltages that are obtained from the input 3-phase voltages using line frequency transformers. Rectifiers with pulse numbers 12, 18, and 24 are common for medium- and high-power applications that require good PF and low THD but do not have stringent constraints on size and weight.

22.6.3 Controlled Thyristor Rectifiers

Diode bridge rectifiers do not have any regulation capability and the output dc voltage varies with changes in line and load. This drawback is overcome by controlled thyristor rectifiers. Thyristor rectifiers are primarily used in medium- to high-power applications where regulation of the output dc voltage is required but line current quality and PF are not important (or can be corrected externally). Increasing concerns for power quality have resulted in reduced applications for these converters. High-power dc motor drives, especially those used in traction, battery chargers, and high-voltage dc (HVDC) transmission are the most common uses for these converters.

To understand the operation of thyristor rectifiers it is first necessary to know the basic terminal characteristics of thyristors. Thyristors, also called silicon-controlled rectifiers (SCRs), are high-power semiconductor devices that can block voltage of either polarity and conduct current in one direction (from anode to cathode). They can be switched on by applying a current pulse to their gate terminal when forward biased (positive voltage from anode to cathode), and can be switched off only by reducing the device current to zero.

Single-Phase Thyristor Recitifier. Figure 22-27*a* shows a single-phase *fully controlled* thyristor rectifier. The output has to be inductive for proper operation. For analysis presented here, it is assumed that I_o is constant and that there is no source impedance. During the positive half of the line cycle ($v_s > 0$), T_1 and T_4 are switched on after a delay angle α from the zero crossing of v_s. The angle α is commonly called the *firing angle*. With T_1 and T_4 on, $i_s = I_o$ and $v_{rect} = v_s$. When v_s reverses in

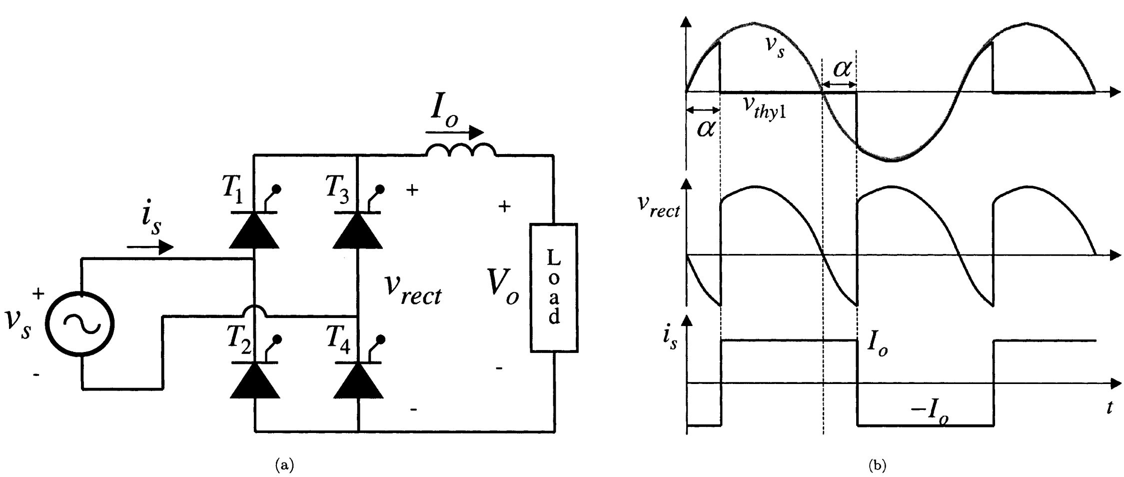


FIGURE 22-27 Single-phase thyristor rectifier: (*a*) circuit, (*b*) waveforms.

polarity thyristors T_1 and T_4 keep conducting since the current through them has not been reduced to zero. During the negative half cycle, T_2 and T_3 are switched on after angle α from the zero crossing of v_s. At this point current is transferred from (T_1, T_4) to (T_2, T_3). In reality there is some finite source inductance, due to which the current transfer takes some time [4]. Once T_2 and T_3 start conducting $i_s = -I_o$, and there is a reverse voltage across T_1 and T_4 that keeps them in the off state. Steady state operating waveforms are shown in Fig. 22-27*b*. The average dc voltage across the load is given by

$$V_o = \bar{v}_{rect} = \frac{1}{\pi}\int_{\alpha}^{\alpha+\pi} \hat{V}_{ph}\sin(\omega t) \quad d(\omega t) = \frac{2\cos(\alpha)}{\pi}\cdot \hat{V}_{ph} \tag{22-53}$$

V_o can be controlled by varying $\alpha \in [0°, 180°]$. It is maximum for $\alpha = 0°$, where the thyristor rectifier behaves exactly like a diode bridge rectifier, and zero for $\alpha = 90°$. For $\alpha > 90°$, $V_o < 0$, and power is transferred from the dc side to the ac side. Where bidirectional power flow is not required, thyristors T_2 and T_4 are replaced by diodes resulting in a *half-controlled* rectifier [43]. Total harmonic distortion of the input current is same as that in the diode bridge rectifier, but the fundamental component of the input current lags the input voltage by angle α, leading to a displacement PF of $\cos(\alpha)$. Thus, regulation of output voltage is achieved at the expense of lower PF.

Three-Phase Thyristor Rectifier. The 3-phase thyristor rectifier is shown in Fig. 22-28*a*. Similar to the single-phase case, each thyristor is switched with a delay of angle α after the anode to cathode voltage across it becomes positive. Each thyristor conducts for 120°, so the input line currents are quasi-square waves with magnitude equal to I_o, as shown in Fig. 22-28*b*. The average output voltage is $V_o = (3\cos(\alpha)/\pi)\hat{V}_{LL}$. For $\alpha > 90°$, power flows from the dc side to the ac side, and the converter acts like an inverter. For unidirectional power flow the three lower thyristors can be replaced with diodes to give a half-controlled version. As with diode bridge rectifiers, thyristor bridges can also be used to obtain 12 (or higher) pulse rectification resulting in lower THD for the input current and reduced size for the output filter. Due to the bidirectional power flow capability of this converter and very high voltage and current rating of thyristors, series connected thyristor rectifiers are utilized in HVDC transmission systems [44].

22.7 AC TO AC CONVERSION

In applications where a controllable 3-phase ac voltage has to be synthesized, the most common strategy is to first rectify line frequency ac to obtain a dc voltage, and then use a 3-phase inverter. The dc link requires a substantial electrolytic capacitor, which filters the dc voltage and also provides energy storage for short duration line voltage sags and interruptions. Capacitors add significant size and cost, and electrolytic capacitors also have the problem of lower reliability. To reduce the number of stages from two to one, and to eliminate the electrolytic capacitor, there has been a significant research effort in direct ac to ac conversion.

Thyristor-based *cycloconverters* [43] have been used extensively for direct ac to ac conversion. In these converters, a low-frequency ac waveform is synthesized by a piecewise combination of the available input ac voltages. These converters have been used for high-power variable frequency ac drives. However, they have limited control over the magnitude, frequency, and quality of the output voltage, and quality of the input line current.

Recently, matrix converters utilizing controllable bidirectional switches and PWM have been developed. As the name suggests, a matrix converter consists of a matrix of switches connecting each input phase to each output phase as shown in Fig. 22-29*a*. All the switches, denoted by square boxes in the figure, need to have bidirectional voltage blocking and current conduction capabilities. So far, a single semiconductor switch with these capabilities has not been invented. Thus, the switch has to be realized using a combination of existing power devices. One implementation is shown in Fig. 22-29*b*.

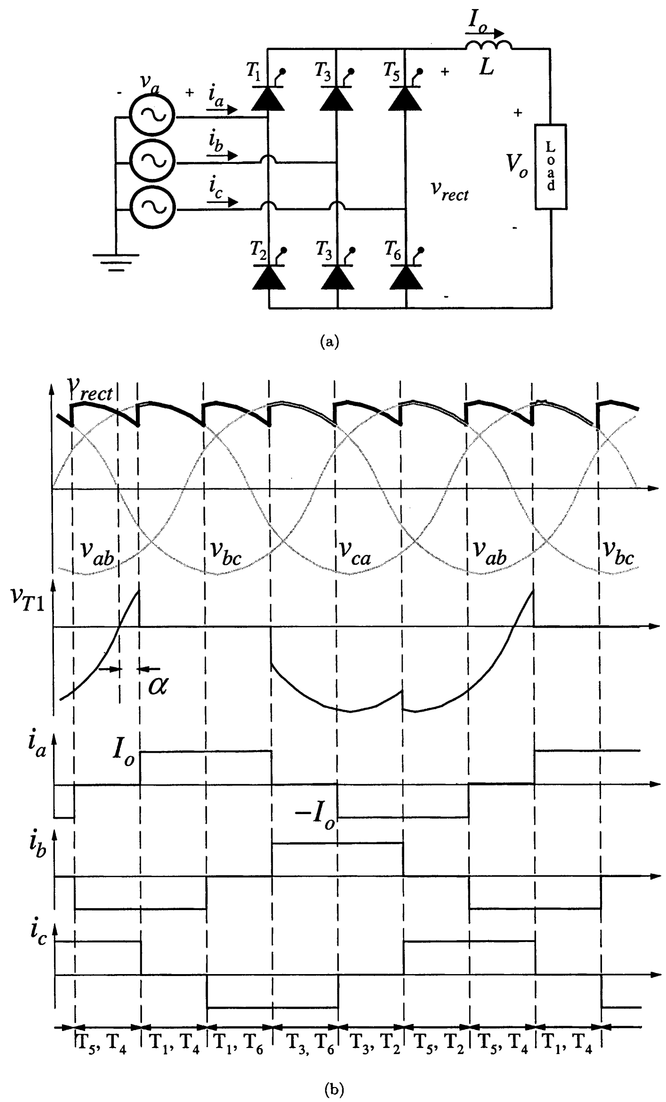


FIGURE 22-28 3-Phase thyristor rectifier: (*a*) circuit, (*b*) waveforms.

At any time instant each of the output phases is connected to one of the input phases, and more than one output phase may be connected to the same input phase. Selecting appropriate switches and using PWM, output voltages with continuously variable amplitude and frequency can be synthesized. The synthesis is most easily understood by means of space vectors [27, 45]. The total number of meaningful switching combinations are 27. Out of these, 6 lead to output voltage space vectors rotating at the input line frequency, 18 lead to stationary output voltage space vectors, while 3 lead to zero vectors. As in the dc to 3-phase ac case, on an average basis, a desired output voltage vector can be synthesized by using a combination of the stationary nonzero and zero space vectors. There is sufficient flexibility to ensure that the input power factor is unity. Considerable research has also

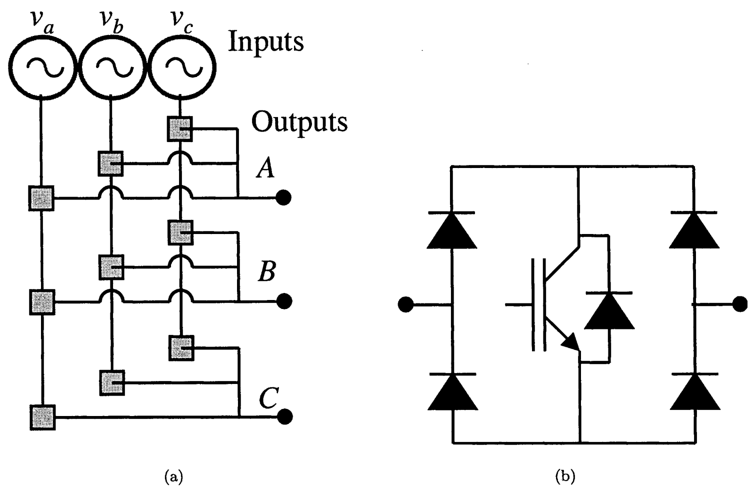


FIGURE 22-29 Matrix converters: (*a*) 3-phase matrix converter, (*b*) bidirectional switch.

been done to ensure proper operation of matrix converters under unbalanced line conditions. However, as shown in Ref. [46] the output voltage of a matrix converter has a theoretical limitation of $V_{LL,op} = (\sqrt{3}/2)V_{LL,in}$, which is considerably lower than that obtainable with an ac-dc-ac configuration ($V_{LL,in}$). In addition, clamping circuits and small input and output filters are required for proper operation.

So far, matrix converters have not been commercially successful. This is chiefly due to the number and cost of bidirectional switches, limitation on the maximum amplitude of the output voltages, and lack of energy storage, which is becoming increasingly important to provide *ride-through* capability during short duration line failures.

22.8 PROBLEMS CAUSED BY POWER ELECTRONIC CONVERTERS AND SOLUTIONS

The two main problems caused by power electronic converters are non-sinusoidal currents injected into or drawn from the utility and conducted and radiated electromagnetic emissions potentially leading to electromagnetic interference (EMI). In addition, power electronic converters have a negative incremental impedance, that is, as the input voltage reduces they draw higher current in order to supply a constant load power. Thus, as the total load supplied by power electronics increases, their effect on power system stability will also increase.

22.8.1 Harmonics and Power Factor Correction

As illustrated in Sec. 22.6, diode bridge rectifiers can inject significant current harmonics, and also result in a lagging power factor for the current. Both harmonics and lagging power factor lead to increased line losses. The harmonic currents in conjunction with the source impedance to the point of common coupling (pcc, where other loads are connected), lead to a distortion in the pcc voltage. Triplen harmonics from single-phase rectifiers contribute to zero-sequence currents when connected in a 3-phase system with a common neutral. If there are several single-phase diode bridge rectifiers connected across different phases and the neutral in a 3-phase system, it can lead to a very significant current flowing in the neutral wire, which is rated to carry only a small current due to load imbalance. Grid connected inverters can also inject currents at the switching frequency unless the switching frequency is sufficiently high or appropriate filters are added.

Due to these concerns, standards have been formulated to limit the amount of distortion in current drawn from and injected into the utility supply. The IEEE 519 [47] specifies the maximum value for individual current harmonics and the overall THD for different applications and power levels. In addition, it also specifies a limit on the distortion in the pcc voltage. However, so far, the IEEE 519 is only a recommendation and not enforced by law. The IEC 1000-3 standard [48] specifies similar limits, and has been made into the EN61000-3-2 European norm. Thus, all electronic equipment sold in Europe has to comply with it. As a result, single-phase and 3-phase unity power factor (UPF) rectification techniques and power factor correction methods have been developed. Unity power factor rectification implies that the rectification technique ensures unity power factor operation. Power factor correction implies that the equipment itself does not draw input current at unity power factor; however, an additional circuit, such as an active, passive, or hybrid filter, is added to ensure that the line current is sinusoidal and in phase with the input voltage. The simplest single- and 3-phase UPF techniques are discussed below. A more comprehensive treatment can be found in Refs. [5, 49].

Single-Phase Boost UPF. The single-phase boost UPF, shown in Fig. 22-30*a*, is the most popular circuit for shaping the input current to be sinusoidal, maintaining it in phase with the input voltage, and regulating the output dc voltage [4, 5].

Basic Operation. The circuit consists of a full-bridge diode rectifier followed by a boost converter. The input voltage to the boost converter, v_{rect}, varies based on the diode conduction; ideally it is a rectified sinusoidal waveform. The output voltage v_o has to be greater than $\max(v_{rect}(t))$ for proper operation; to work with both 120 and 220 V ac inputs (also called universal input voltage range) the output is usually chosen in the range of 380 to 400 V. The boost converter operates so that there is a quasi-steady condition at each point of the input sine wave. Thus, the duty ratio varies as

$$d(t) = 1 - \frac{v_{rect}(t)}{v_o} = 1 - \frac{\hat{V}_s\left|\sin(\omega t)\right|}{v_o} \tag{22-54}$$

The duty ratio is controlled to satisfy two objectives: regulation of output voltage and shaping the inductor current to follow the waveshape of $v_{rect}(t)$. This is done by using the cascaded control structure shown in Fig. 22-30*b*. The error compensator for the outer voltage loop, $G_v(s)$, generates a reference signal for the amplitude of the inductor current (or the rectified input current), $\hat{I}_L$,*ref*. The effect of input voltage magnitude is fed forward by dividing the output of $G_v(s)$ by a signal proportional to square of the rms input voltage. The amplitude reference $\hat{I}_{L,ref}$ is multiplied by the waveshape of the rectified voltage to obtain $i_{L,ref}(t)$, the reference signal for the inductor current. The inductor current $i_L(t)$ is controlled to $i_{L,ref}(t)$ by using average current mode control. Steady-state waveforms are shown in Fig. 22.30*c*.

Control Issues. The current control loop has a bandwidth of the order of 10 kHz, required to follow the rectified sine wave shape accurately enough so that input current has acceptable THD. After each zero crossing of the input voltage, the positive rate of change of inductor current required is very high, while the voltage to effect this change, $v_{rect}(t)$, is close to zero. Thus, there may be a significant error in the current during this time, and the resulting distortion in the input current is called *cusp distortion*. Cusp distortion contributes to harmonics and lagging power factor.

The output voltage necessarily has a 120 Hz ripple component due to variation of power in single-phase systems. The error in the voltage feedback loop also has this ripple component. If this component is compensated by the control loop, it leads to a 120 Hz component in the output of $G_v(s)$, which after multiplication with the rectified voltage waveshape leads to a third harmonic component in the input current. To avoid this, the bandwidth of the voltage control loop is intentionally kept very low—in the range of 5 to 20 Hz. The converter, therefore, has very poor dynamic response with respect to changes in load. The input voltage feed forward signal also has to be filtered heavily to attenuate the ripple component sufficiently. Thus, although the feed forward improves the dynamic performance with respect to line changes, it is still quite slow. Further, to keep the voltage control design simple, a pure integrator is not added in the voltage

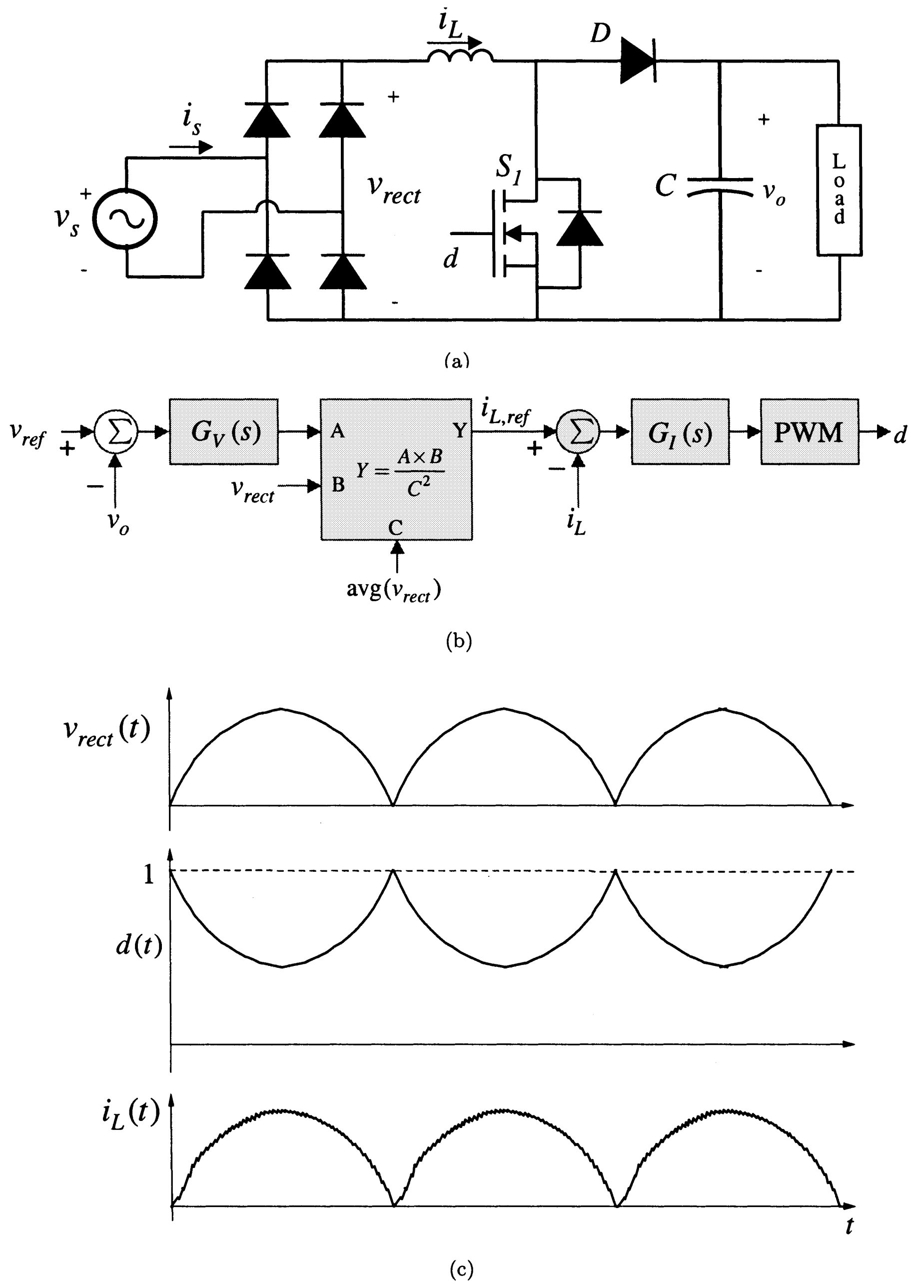


FIGURE 22-30 Single-phase boost UPF (*a*) circuit, (*b*) control block diagram, (*c*) steady-state waveforms.

control loop leading to a nonzero steady-state error. Typically, the output voltage drops by about 10 to 20 V from no load to full load.

The entire control, including the multiplier and divider circuit, can be implemented by means of a single analog integrated circuit (IC) such as the UC3854 [9], although digital implementations are also being explored. The availability of a single chip controller, easy to design power stage, and standardized control design equations (e.g., see the UC3854 application note on TI's website [9]) has made this topology very popular for single-phase applications. However, there are several other topologies that may be suitable for specific applications [5, 49].

The dynamic performance of the single-phase boost UPF is not adequate if the output is directly connected to a load. Usually, isolation of output from input line is a requirement, and the required output voltage may be significantly different from 400 V. Thus, an isolated dc-dc converter stage is required at the output of this UPF. This additional converter takes care of regulation, isolation, and scaling the voltage.

3-Phase UPF. In the 3-phase case, the diode bridge rectifier can be replaced by a 3-phase PWM inverter with external inductances in series with each phase. By appropriate control, this converter can transfer power from the ac side to the dc side. Input currents can be controlled in a manner similar to the single-phase case—waveshape determined by input voltages and amplitude determined by the voltage control loop. The converter operates in boost mode so the output dc voltage is regulated to a value higher than the peak of the input line-to-line voltage. In this case, dynamic response of the voltage control loop is limited by the nonminimum phase nature of the system. This 3-phase rectifier is very expensive compared to its diode bridge counterpart, and its cost is justified only if bidirectional power transfer capability is required. Thus, several alternate solutions for 3-phase power factor correction have been proposed [5, 49].

22.8.2 Electromagnetic Interference

Electromagnetic interference refers to maloperation of electronic equipment due to electromagnetic emissions which originate in the "surroundings". Here, surroundings refer to other electronic equipment that are electrically connected or spatially close. Electromagnetic emissions emanate from transitions in currents and voltages, especially fast transients involving high magnitudes. Thus, power electronic converters can be major source of emissions leading to EMI. Within a converter the main sources are as follows:

- Current transfer between rectifier diodes.
- Switching of semiconductor devices: capacitive coupling between the high voltage switching terminals of the device and heat sink (which is usually grounded) leads to high-frequency common mode and differential mode currents.
- High-frequency inductors and transformers.
- Long wires or printed circuit board (PCB) traces carrying high-frequency currents.

Emissions can be classified into two types: *conducted* and *radiated*. In the conducted case, high-frequency currents produced due to switching transients flow through the parasitics in the power converter, and are manifested as common mode and differential mode currents at the inputs and outputs of the converter. At the output it can cause EMI problems for sensitive loads, and at the input it creates problems for other equipment connected to the utility. To restrict conducted emissions, EMI filters, good circuit board design, and packaging are needed [7, 50]. *Common mode chokes*, which offer high impedance to common mode currents and very low impedance to differential mode currents, are used to limit common mode currents. Differential mode high-frequency currents are attenuated by using capacitors to earth ground; the capacitance values are limited by the maximum allowable ground current specified by safety regulations like UL478 and UL1283. Complete EMI filters that provide both common and differential mode attenuation are commercially available as packaged units from several manufacturers.

Radiated emissions affect high impedance voltage sensitive circuits (e.g., MOS input circuits) that are in close proximity to the converter. To reduce radiated emissions, inductors and transformers are shielded (especially around air gaps), a careful PCB design that minimizes long paths for high frequency and switched currents is carried out, and long wires carrying high-frequency currents are twisted with their return lines. "Twisting" can also be approximately achieved for PCB traces. In addition, the entire converter may be shielded by a grounded metal enclosure.

Depending on the application power converters have to comply with an emissions standard. Emissions standards, formulated by the FCC [51], VDE [52], and the military, dictate the maximum amount of conducted and radiated emissions allowed for different kinds of electronic equipment.

22.9 APPLICATIONS OF POWER ELECTRONIC CONVERTERS

The foregoing sections described the principles and control of different power converters. This section describes the requirements of specific applications, and how power electronic converters are utilized for these. The next section is devoted to utility applications of power electronics.

22.9.1 DC Power Supplies

DC power supplies are required for powering different components in electronic equipment. Their specific features and level of sophistication depend on the application. For example, in a *bias supply*, required for various analog components in a power converter, the main requirements are isolation and a (relatively) large tolerance in the output voltage level. The input to the bias supply may be derived from a regulated dc bus, so line regulation is not required. A simple push-pull converter without an output inductor can be used for this purpose. For use inside a sensitive instrument, the power supply requirements would include strignent regulation with respect to line and load, good dynamic response, compliance with EMI and input power quality standards, and adequate energy storage for normal operation during short duration line failures. Features commonly required in dc power supplies are as follows.

Output Voltage Regulation with Respect to Line and Load. For high-current low-voltage power supplies there may be an additional requirement of regulating the voltage at the load terminals. This requirement, called remote sensing, accounts for voltage drops in the connecting wires. Dynamic requirements of response time and overshoot/undershoot in the output to step change in load current are also specified.

Output Current Limit. This may be a fixed value, or have a *foldback* characteristic [6]. With the latter, once the output current exceeds a certain value, the current limit is varied as a function of output voltage—decreasing as the output voltage decreases.

Isolation. Electrical isolation between output and input.

Soft Start. This is required to limit the inrush current for initial charging of capacitors.

Holdup Time. For line powered applications, there is usually a requirement for holdup time, time for which the power supply should operate normally in the absence of the input voltage. Usually energy storage is provided by adding capacitors; sometimes auxiliary capacitors charged to a higher voltage to store more energy are used.

Sleep Mode. Typically power supplies have a low efficiency at light load. Thus, in battery-powered and portable applications light load condition is detected and the power supply operation is changed to reduce its losses, for example, by reducing the switching frequency.

Power Factor Corrected Front End. For single-phase applications, it also helps in operation with universal input voltage range (100–240V, 50/60 Hz).

EMI Compliance. With a specified standard.

Environmental compliance. For temperature, humidity, and altitude of operation. Some features can be implemented with standard PWM control ICs leading to reduced number of components and design simplification. Commercial availability of low power dc-dc converters modules with standard input and output voltages has increased significantly in recent years. Thus, a custom power converter design may not be required, unless there are special requirements as in space, defense, and some instrumentation applications.

A specific application of dc power supplies is in digital systems. All digital systems need a power electronic converter to provide the requisite supply voltages. With the digital supply voltages going down and the clock frequencies going up, the requirements are increasingly toward low voltages and high currents. For a high-speed microprocessor, the required values are around 2 V and a few hundred amperes. The dynamic requirements are also very demanding since the microprocessor load can go from almost zero current to full load current in a few microseconds. Furthermore, digital components operating at different supply voltage levels may be used in a complete digital system. For these applications, the intermediate bus architecture is used. First, a 12 V (or a lower voltage) bus is derived from the input supply using an intermediate bus converter (IBC); the IBC may also provide isolation from the input. This low-voltage bus is then input to several point-of-load (POL) converters, which are located close to the loads and provide the required steady state and dynamic voltage regulation. Point-of-load converters do not provide isolation between input and output. Intermediate bus converter and POLs with standard input and output voltages are commercially available as modules from several manufacturers (e.g., see Refs. [17, 53]).

22.9.2 Electric Drives

Introduction. Traditionally electric motors have been powered by direct connection to ac line, or to dc voltage obtained from a rectifier. However, this usually results in inefficient operation due to lack of control. Electric motors powered by appropriately controlled power electronic converters lead to significant increase in the overall system efficiency due to the advantages of variable speed operation [54]. Electric motor load has more than half the share of electric power consumption in the United States, of which about half is in industrial applications. Thus, any increase in system efficiency due to electronic controlled motor drives can lead to large savings for the company and overall electricity consumption. Moreover, performance advantages of fast dynamic response and very accurate control over speed and position are obtained.

Squirrel cage induction machines (IMs) are the most widely used electric motors. This is due to the advantages of simple, robust, and low-cost construction, ease of powering, and lower maintainance requirement due to the absence of brushes. Thus, control of induction machines for better performance and improved efficiency has been researched extensively. The invention of vector control [33], and direct torque control [34, 35] have made the dynamic performance of induction machines similar to that of a dc machine. The popularity of vector control was also aided by development of PWM inverters and space vector modulation described in Sec. 22.5.3. Developments in permanent magnets and power electronics, and increasing concerns for size, weight, and efficiency have led to significant interest in permanent magnet synchronous machines (PMSM) and brush-less dc (BLDC) machines. Permanent magnet synchronous machines and BLDC motors have the highest operating efficiency and the highest power density of all the motors [55]. Further, due to development of power electronics, the switched reluctance machine (SRM), the first electromechanical machine invented in the early 1800s, has also been commercialized [56, 57]. This machine has a simple, robust, and inexpensive construction, and is powered from an inherently robust power converter topology. Switched reluctance machines are most popular in the low-to-medium power range and high-speed applications requiring medium performance. Squirrel cage induction, PMSM, Synchronous Reluctance, and BLDC drives require a 3-phase inverter for powering the stator windings and no electrical excitation for the rotor. Switched reluctance machines have a special power converter that has independent sections for individual phases. For illustration, a vector controlled ac drive and an SRM drive are discussed here. Details of these drives may be found in Refs. [28, 30, 36, 37, 54–57].

Vector Controlled AC Drive. Figure 22-31 shows the basic block diagram of a vector-controlled PMSM drive. The structure is identical for the squirrel cage induction motor drive. The input stage is a single- or 3-phase diode bridge rectifier, the output of which is filtered to obtain a dc. The input current during initial charging of the dc bus capacitor is limited using either (1) parallel combination of a controllable switch and a current limiting resistor connected in series with the main power path, or (2) a half-controlled thyristor rectifier. The dc voltage is converted to variable frequency variable magnitude 3-phase ac by means of a 3-phase PWM inverter. During *dynamic braking*, energy stored in the inertia of the machine and load is transferred back to the electrical circuit. If the rectifier has bidirectional power transfer capability, this energy can be transferred to the source. Otherwise, the energy charges the dc bus capacitance. To limit the dc bus voltage, a brake resistor is connected in parallel with the dc bus by means of a controlled switch.

The switching frequency of the inverter in ac drives is usually between 20 and 30 kHz, since the machine inductance is usually sufficient to filter the output current. Thus, IGBTs are commonly used for this application. MOSFETs are used if the dc voltage is below 600 V and a higher switching frequency is required in order to reduce the output current ripple. Intelligent power modules (IPMs) containing all the power semiconductor components along with gate drive and protection circuits in a single package are now available from several manufacturers (e.g., Refs. [58–62]). Use of IPMs results in considerable design simplification, size reduction, and smaller design cycles.

The drive control is carried out using a microcontroller or digital signal processor (DSP), with the use of DSPs becoming more popular due to increased computational requirements. In an industrial application, the DSP controls the machine in response to command signals (for speed or position) obtained over a communication interface bus. AC machines are usually controlled using either vectrol control or direct torque control. From the power electronics point of view, this requires control of the output currents with a high control bandwidth (of the order of 1 kHz). This is done in a cascade manner: the current control algorithm generates reference voltage vectors, which are then synthesized using space vector PWM. Both current control and PWM are implemented on the DSP. The final output of the DSP are switching signals for each of the three inverter legs and the brake switch. These signals are optically isolated and input to the gate drive circuitry for each of the

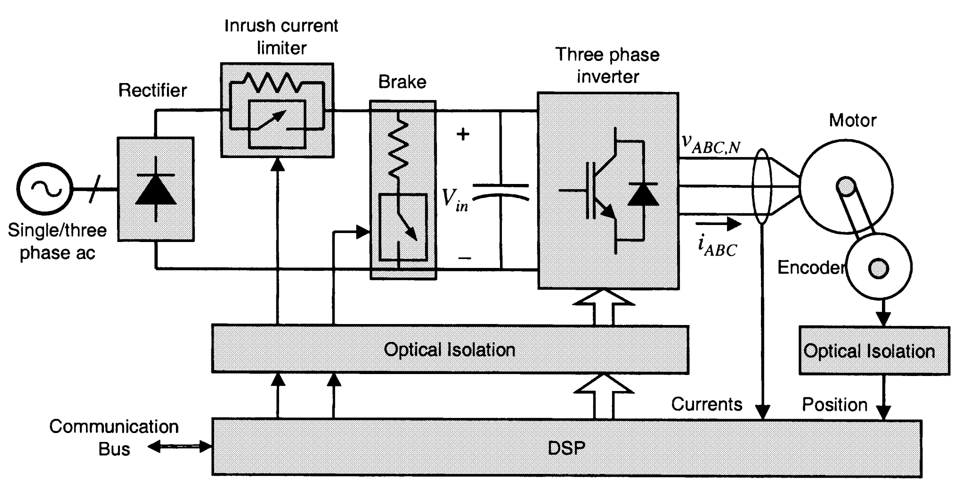


FIGURE 22-31 Vector controlled 3-phase ac drive.

controlled switches. Output currents of the inverter are sensed using either hall effect sensors or sense resistors. The analog current signal is converted to digital form using the internal analog to digital converters (ADCs) of the DSP. The dc bus voltage may also be sensed for feedforward of input line variations and dynamic braking. An optical encoder is used if direct rotor position sensing is required. Since the encoder outputs are already in digital form they can be easily interfaced. However, due to close proximity of the encoder and the machine windings, which are subject to pulsed voltages generated by the inverter, these signals have very high common mode noise. To solve this problem, differential mode line drivers and receivers, and optical isolation of encoder signals are commonly employed.

SRM Drive. The basic structure of an SRM drive is similar to an ac motor drive. However, the power converter is quite different owing to the machine characteristics. Figure 22-32*a* shows the cross-section of a 4-phase SRM. The SRM has saliency on the stator as well as the rotor. Each phase consists of concentrated coils wound on diametrically opposite stator poles. If phase *a* carries current i_a then the closest set of rotor poles are attracted to the stator phase *a* poles. Once the rotor poles are aligned with the stator poles, there is no torque on the rotor and i_a has to be reduced to zero. By energizing and deenergizing phases in sequence a continuous rotation can be achieved.

To build up and reduce winding current, the converter needs to have bidirectional voltage capability. However, since the direction of torque is independent of the current polarity, only unidirectional current capability is required. At low speeds, the back emf of the machine is very low so the winding current has to be controlled usually by hysteretic current control. This is usually done by letting the current freewheel with a zero voltage across the winding terminals, commonly called application of a zero voltage loop (ZVL). To determine which phase needs to be energized or deenergized rotor position information is required, this may be sensed or estimated, and discrete position information is sometimes sufficient [57]. Usually the total torque produced by the machine has significant ripple, and unless instantaneous torque control is a requirement, average torque control is implemented. The average torque depends on the turn-on and turn-off angles (rotor positions with respect to the phase winding where that phase is energized and deenergized, respectively), and the current magnitude if current control is used [56]. The control is usually implemented based on results of numerical simulation and then tuned experimentally.

Ray and Davis published one of the earliest paper on power converters for SRM [63], and suggested the asymmetric bridge converter, which is also known as the classical converter. The circuit topology, shown in Fig. 22-32*b*, is similar to a two-switch forward converter [4] and fulfils all the

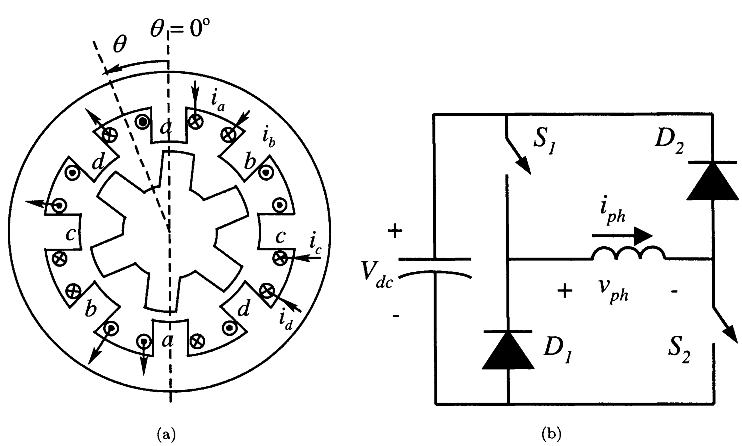


FIGURE 22-32 SRM Drives: (*a*) Cross-section of a 4-phase 8/6 SRM, (*b*) asymmetric bridge converter.

basic requirements. Only one phase is shown, the circuit being identical for the other phases with a common connection to the input dc bus. The basic operating modes of the converter are as follows:

Energization	(S_1, S_2) on, (D_1, D_2) off	$\upsilon_{ph} = V_{dc}$
Deenergization	(D_1, D_2) on, (S_1, S_2) off	$\upsilon_{ph} = -V_{dc}$
ZVL	(S_1, D_2) on, (S_2, D_1) off, or	$\upsilon_{ph} = 0$
	(S_2, D_1) on, (S_1, D_2) off	

where all switches and diodes have been assumed ideal and the winding resistance has been neglected. This converter is extremely reliable and fault tolerant. The problem of "shoot-through" due to spurious turn-on of the two controlled switches in the same leg is avoided, since the phase winding is in series with the controlled switches. Shoot-through can be a major problem in full-bridge dc-dc converters and 3-phase inverters. Although this converter topology has become very popular and a packaged module similar to 3-phase ac inverters is now available, there are several other converter topologies, which may be suitable for specific applications. A detailed comparison of SRM converter topologies can be found in Refs. [56, 64–66].

DC motor drives use a dc-dc converter—buck, buck-boost, or full bridge depending on the requirements—and armature current control for controlling the instantaneous torque. The field current may also be controlled to reduce back emf at high speeds. Their popularity has been decreasing due to maintenance requirements and speed limitations of brushes. As the cost of power electronics reduces, better and low-cost permament magnets are developed, better control and estimation techniques are implemented on lower cost DSPs, and energy efficiency gains priority, it is expected that PMSM and BLDC drives will increase their share in motor drives. Switched reluctance machines are also expected to gain ground in high speed and harsh environment applications where energy efficiency is not very important.

22.9.3 Battery Charging

Battery charging is a very large part of the ac-dc and dc-dc converter applications. The end use for batteries is in telecommunications, electric and hybrid electric vehicles, portable electronics, and energy storage for improving power system stability. For a lead acid battery, there are three standard operating modes for the battery charger: (1) constant current (or bulk) charging during low state of charge, (2) constant voltage charging after about 80% state of charge, and (3) "float" or trickle charge after the battery reaches its open circuit voltage. Thus, the requirement on the power converter is to operate at its maximum current rating over a significant output voltage range with high efficiency. In addition, isolation from input line and power factor correction are usually required.

In recent years, lithium-ion (Li-ion) batteries have gained widespread usage for portable applications like laptop computers, PDAs, and cell phones. They have also been proposed for automotive applications, but so far nickel metal hydride (NiMH) batteries are used in commercially available hybrid electric vehicles (HEV). The advantages of Li-ion batteries are high energy density, higher cell voltage of about 4.2 V, and low self-discharge rate [67]. The disadvantages are high sensitivity to electrical stress and limited temperature range. The cell voltage has to be monitored for overvoltage (which is very close to the nominal open circuit voltage) to avoid catastrophic failure, and undervoltage to maintain battery life. Linear chargers are used for single-cell batteries requiring low charging current, while switch mode converters are used for high-voltage and high-current charging. The charging technique is similar to lead acid—constant current followed by constant voltage. However, pulse charging (alongwith periodic discharge and periodic relaxation pulses) have also been proposed for improving battery life. A further complication with Li-ion is that the dependence of state of charge on open circuit cell voltage is only on a small voltage window (3.2–3.8 V). Thus, in the absence of a detailed cell model, a columb counting technique, which measures the charging and discharging current and estimates the self-discharge rate, is normally used to determine the battery state.

22.9.4 Fluorescent Lamps and Solid State Lighting

About 30% of the total electricity generated in the United States is consumed for lighting [68]. Thus, there is significant interest in increasing energy efficiency of lighting mechanisms. Incandescent lighting, although the most inefficient, is still the cheapest in terms of $/lumens. However, compact fluorescent lamps (CFLs) have gained ground due to their energy efficiency and are now commercially available as replacements for incandescent bulbs. Besides lighting, fluorescent lamps that emit ultraviolet light are used in several industrial applications like sterilization and curing (drying of coatings).

Compact fluorescent lamps and industrial lamps are usually powered using a switched mode converter called an electronic ballast. In one implementation, the converter consists of a full- or half-bridge inverter with its output connected to an L-C-C (all in series) tank circuit. The dc voltage input to the inverter may be obtained using either a diode bridge rectifier or a power factor corrected front-end. The lamp filaments are connected in series with the L-C-C tank, with the two filaments being in parallel with one of the capacitors. The other capacitor provides dc blocking and also tunes the tank operation to a desirable characteristic. When the lamp is off (gas inside not ionized), it acts as an open circuit, so the inverter is effectively connected to a series resonant L-C circuit. To ionize the gas inside the lamp, the filaments have to be heated sufficiently followed by application of a high voltage (~kV) across the lamp for a short time. Once the lamp is on (gas ionized) it acts a resistive load, and the inverter load is then a series parallel tank circuit—series L-C-C with the equivalent resistance of the lamp in parallel with one capacitor. The inverter outputs a high-frequency square wave, typically in the range of 20 to 100 kHz; lamp operation at these frequencies results in increased light output. The switching frequency is varied to control the ballast operation. Initially, the tank circuit and the lamp are excited at a frequency significantly higher than the resonance frequency. During this time, called the preheat time, the lamp filaments increase in temperature. After a suitable delay, the switching frequency is reduced toward resonance so that the voltage across the lamp increases until the gas inside the lamp ionises and the lamp “ignites”. After ignition, the lamp acts like a resistive load, and the lamp output power is controlled indirectly by regulating the phase difference between the inverter current and voltage. The tank circuit also ensures resonant transitions of the converter switches [69] leading to reduced losses. Several ICs that implement full control of a fluorescent lamp including preheat, ignition, and dimming control are commercially available (e.g., see Ref. [62]).

High intensity discharge (HID) lamps have a large share of commercial lighting such as street lighting, sports facilities, etc. They are of three types—mercury vapor, sodium vapor, and metal halide. The ballast requirements of these lamps differ from those of fluorescent lamps. Operation of these lamps at higher frequencies does not improve light output. Further, they exhibit acoustic resonance in the 10 to 100 kHz range. Thus, so far, electronic ballasts for these lamps have not become very popular.

Power supplies for plasma cutting have requirements similar to fluorescent lamp ballasts. Initially, a high voltage (~10 kV) is applied to produce an arc between two high-voltage electrodes. This is followed by *arc transfer*, transfer of the conducting ionized gas between the electrodes to the space between the negative electrode and the metal being cut (workpiece). Once the arc is transfered the cutting procedure requires a current controlled supply (at lower voltage) connected across the negative electrode and the workpiece. Since there are three terminals, two separate power supplies can be used—one for generating the initial high voltage and the other for the actual cutting operation.

The U.S. Department of Energy has identified solid-state lighting as a means to increase lighting efficiencies in the near future [68]. Recent and expected breakthroughs in semiconductor and organic light emitting diodes (LEDs) are the main impetus behind this. Another reason promoting the use of LEDs is the European Union’s directive on removal of hazardous substances (RoHS) in electrical and electronic equipment (Directive 2002/95/EC). The directive aims at phasing out the use of mercury and lead (among other substances), which are chief components of compact fluorescent and HID lamps. LEDs do not use either of these materials. Although the RoHS directive, going into effect in July 2006, makes exceptions for fluorescent and some HID lamps used for lighting, it is expected that these will be covered in future revisions of the directive. Light emitting diodes are already being used for traffic lights, exit signs, and as indicator lights in automobiles. They have the advantages of faster turn-on time, increased efficiency, vibration and shock resistance (required for automotive), and longer operating life. For solid-state lighting, light output of several series connected LEDs, called *strings*, is used. A switched mode converter with a dc current regulated output is used to power the LED string. The output current is regulated to provide dimming control and to avoid failure due to overcurrent [70]. When LEDs start replacing HID lamps, incandescent bulbs, and CFCs for commercial/residential lighting, they will open up a huge market for efficient and low-cost power electronic converters.

22.9.5 Automotive Applications

In recent years, electronically controlled load in automobiles has increased significantly. Further increase is expected due to more comfort features, and potential replacement of some mechanical systems by all electrical systems like active suspension and power steering. The present bus voltage in automotive systems (14 V), decided by the alternator charging voltage, implies very high currents for the expected high power consumption. To keep the current levels manageable, the 42 V power net has been proposed, and adopted by most auto companies and suppliers for future generation automobiles [1]. The choice of 42 V has been made on the basis of safety considerations, load dump overvoltage transients, and optimal utilization of silicon in power semiconductor devices. It is expected that there will be a dual voltage system, with both 42 and 14 V loads, although the exact configuration has not been standardized. Thus, a dc-dc converter, possibly bidirectional, will be required for interconnection between the two system voltages. Furthermore, converters with sophisticated controls will be required for features like active suspension and power steering [71].

Hybrid electric vehicles (HEVs) are now gaining commercial success due to their higher mileage, lower emissions, and tax breaks. The basic ideas in an HEV are:

- Operate engine at optimal efficiency, achieved at high speeds.
- During initial acceleration and at low speeds, power is supplied by a battery powered motor leading to reduced idling losses and emissions.
- For high acceleration, power is derived from both the engine and the battery.
- During high speeds, the battery is charged from a generator connected to the engine. The motor used for generating mechanical power can also be used as a generator.
- During braking, the generator feeds energy back to the battery.

Power electronic converters are required to control the motor (and generator for a separate machine), and for battery charging. In HEVs, the battery has to be maintained at less than full charge to receive regenerative energy during braking. Hybrid and completely electric vehicles with fuel cells as the power source and some means of energy storage are also being developed by several auto manufacturers. These require power converters to interface fuel cell, energy storage element, motor/generator, and other electronically controlled loads.

22.10 UTILITY APPLICATIONS OF POWER ELECTRONICS

22.10.1 Introduction

Power electronics has the potential to change the landscape of power generation, transmission, distribution, and end use. Growing energy demand, coupled with economic, environmental, and political restrictions on newer generation and transmission infrastructure, means that the existing resources operate near their stability limits. As a result dynamic instability, inter-area oscillations, voltage instability cascading to major blackouts have become issues of real concerns today. Power electronic technology is widely considered to be one of the key components of the grid modernization.

Distributed generation is an important energy option for the twenty-first century and is a key element of restructuring of the electric utility. Distributed generation using renewable sources such as solar, wind and tidal energy, and hydrogen, and their interface through power electronic converters, represents one of the most promising paths to sustainable energy. Also, several modern loads such as the processing plants in semiconductor industry or data centers require clean and uninterrupted power. Power electronic-based power quality solutions are essential for these loads to mitigate problems such as voltage sags, harmonics, and flicker in line voltage.

Widespread use of power electronics in power systems is further fueled by dramatic advances in power semiconductor materials and devices, especially those based on silicon carbide (SiC) [72, 73], and advances in the fields of wide area power system monitoring and communication. This section briefly describes the major power electronics applications in power systems, namely, flexible ac transmission systems (FACTS), custom power, and interfacing distributed generation with electric grid. Figure 22-33 shows the interconnected power system network with some of the major power electronics highlighted.

22.10.2 Flexible AC Transmission Systems

Flexible ac transmission systems is a collective term for different types of power electronic devices/converters based systems that are capable of controlling power flow in high-voltage ac transmission systems [74, 75]. With advances in power semiconductor devices, PWM methods, and control theory, the use of FACTS devices has seen a significant increase. Flexible ac transmission systems devices are capable of the following major functions:

- Control power flow along desired transmission corridors, which is critical for a deregulated utility; they can also minimize loop flows.
- Increase transmission capacity without requiring new transmission infrastructure.
- Improve transient, dynamic, and voltage stability, and provide damping for inter area oscillations.

Different types of FACTS devices control different parameters of the transmission system like the effective line impedance, bus voltage magnitudes, or phase angles, to control power flow, and to

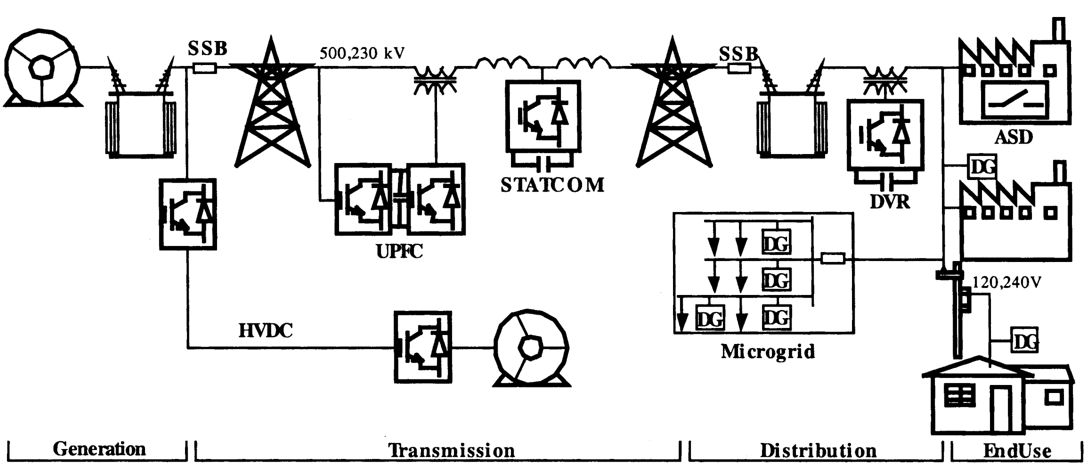


FIGURE 22-33 Power electronics applications in power systems.

increase stability margins. Consider the single line diagram of a 2-bus system shown in Fig. 22-34. The real power flow in the connecting transmission line is given by

$$P = \frac{V_1 V_2}{X} \sin(\delta) \tag{22-55}$$

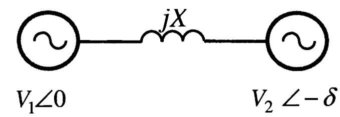


FIGURE 22-34 Single line diagram of a two-bus system.

where V_1, V_2 are the magnitudes of sending and receiving end voltages, respectively, δ is the phase angle between the two voltages, and X is the series line impedance [76]. Flexible ac transmission systems devices control one or more of these three parameters to control power flow and improve stability.

Transient Stability. Flexible ac transmission systems devices have the ability to enhance both transient and dynamic stability of power system networks, thereby enabling increased power flow through existing transmission lines. Transient instability occurs when a major disturbance like fault, line outage, or loss of generation results in large rotor angle deviations leading to loss of synchronism. The rotor angle deviation is governed by the swing equation given in Eq. (22-56).

$$\frac{2H}{\omega_o}\frac{d^2\delta}{dt^2} = P_m - P_e \tag{22-56}$$

where H is the inertia constant in MJ/MVA, ω_o is the synchronous speed, while P_m and P_e are the mechanical power input and electrical power output, respectively. During a fault, the electrical energy drawn from the generator reduces significantly, while mechanical power input remains roughly constant, leading to increasing rotor angle. If the fault is not cleared before a critical time, transient instability occurs. The critical clearing time depends on the electrical power output during the fault and immediately after fault clearance. Since power flow can be controlled continuously using FACTS devices, power during and after a fault can be controlled to improve the stability margin of the system.

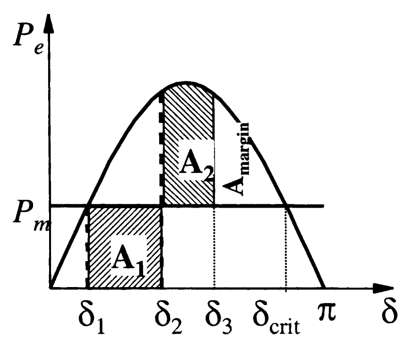


FIGURE 22-35 Equal area criteria for transient stability.

Transient instability is often studied using the equal area criterion as shown in Fig. 22-35 [76]. Initially the mechanical input power input is equal to the electrical power transmitted at an angle δ_1. A fault at the generator makes the electrical power zero while the mechanical input power remains the same, leading to increase in rotor angle from δ_1 to δ_2, at which point the fault is cleared. During this interval, the stored kinetic energy in the machine increases, and increase in kinetic energy is equal to the area A_1 in Fig. 22-35. After the fault is cleared the electrical power transmitted is higher (due to increased phase angle) than the mechanical power input. Hence, the machine begins to decelerate. However, the phase angle increases further due to the stored kinetic energy. The maximum angle is reached at δ_3, when the decelerating energy represented by the area A_2 becomes equal to the accelerating area A_1. If the phase angle extends beyond δ_{crit}, then the system is unstable since decelerating energy cannot balance the accelerating energy. The area A_{margin} between δ_3 and δ_{crit}, represents the transient stability margin of the system. Flexible ac transmission systems devices can improve the margin by dynamically changing the $P - \delta$ characteristics of the system.

Thyristor Controlled Series Capacitor. The earlier FACTS devices were predominantly thyristor based, like the thyristor–controlled series capacitor (TCSC) and static VAR compensator (SVC) [75]. Thyristor controlled series capacitor is a series-connected FACTS device that controls the effective impedance

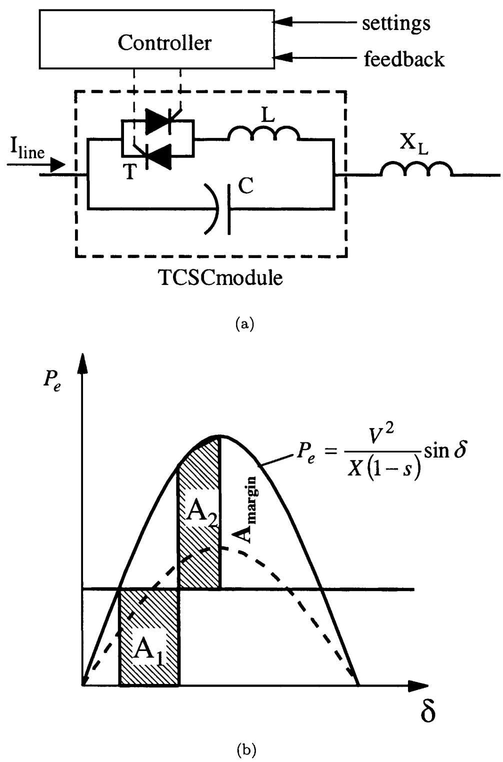


FIGURE 22-36 TCSC: (*a*) schematic (*b*) enhacement of transient stability.

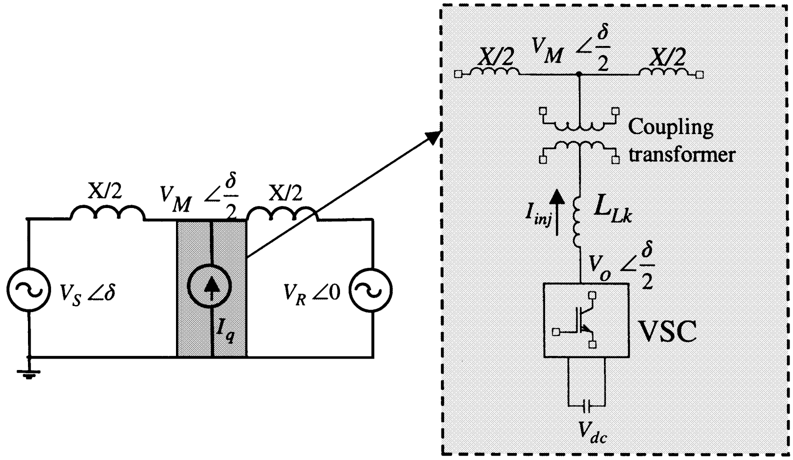
of the transmission line. Figure 22-36*a* shows the basic schematic of a TCSC, which consists of a capacitor in parallel with a thyristor-controlled reactor (TCR). Thyristor controlled reactor is a series combination of an inductor and a pair of phase-controlled thyristors. By suitably controlling the firing angle of the thyristors, the reactance (inductive) of TCR, and therefore, the effective fundamental impedance of TCSC can be controlled continuously. The relationship between the firing angle and the effective TCSC impedance is highly nonlinear. The firing angle α is measured from the zero crossing of the capacitor voltage. As an example, corresponding to an installed capacitive impedance of 0.5 pu (per unit) and inductive reactance of 0.1667 pu, the effective impedance of TCSC can be controlled from about 4 pu capacitive to 2 pu inductive. The effect of TCSC control on the transient stability margin is illustrated in Fig. 22-36*b*. In the figure, *s* indicates the degree of compensation. One of the major advantages of TCSC, when compared with uncontrolled series compensation, is the ability to mitigate sub synchronous resonance (SSR) [75].

Voltage Source Converter Based FACTS. The newer FACTS devices are based on voltage source converters (VSC) implemented using fully controllable devices such as GTO, MCT, IGCT, and IGBT [77]. Within their voltage and current ratings, the VSC-based FACTS devices are capable of injecting any suitable, controlled voltages and/or currents at the line frequency. The main advantages of these FACTS devices, compared to the thyristor-based devices, are the speed of response and the extended control range, which is mostly independent of the line operating conditions. The main VSC-based FACTS devices are the static compensator (STATCOM), the static synchronous series compensator (SSSC), and the unified power flow controller (UPFC).

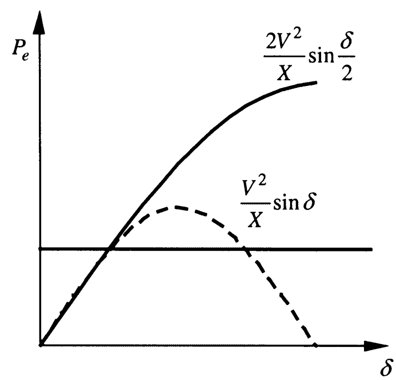
Static Synchronous Compensator (STATCOM). STATCOM is a shunt FACTS device capable of injecting controlled currents at the point of connection with the transmission system [75, 78]. The injected current is usually in phase quadrature (leading or lagging) with the line voltage, so that only reactive power is supplied or consumed by the STATCOM. If real power capability is present, through the use of active energy sources or large energy storage systems, then the injected current can have different phase relationships with the line voltage, thereby extending its control range.

Figure 22-37*a* shows the schematic of a STATCOM connected at the midpoint of a 2-bus transmission system model. The voltage source converter is capable of generating the required fundamental voltage such that the current injected into the system has the desired phase and magnitude to control power flow. The voltage at the dc link is kept constant by large capacitor banks. Losses in the system are compensated, and the capacitor voltage maintained, by drawing a small real power from the transmission system. The voltage source converter is connected to the transmission line through a line frequency coupling transformer, which enables the STATCOM to work with lower voltage switches.

The output voltage of STATCOM (neglecting losses) is controlled to be in phase with the line voltage. Hence, the system can be modeled as two in-phase, line frequency voltage sources,



(a)



(b)

FIGURE 22-37 STATCOM: (*a*) midpoint connection, (*b*) variation of power with phase angle δ.

connected by a reactor (usually the leakage inductance of the coupling transformer), which results in the current I_{inj} being purely reactive. Referring to Fig. 22-37*a*, if the magnitude of V_o is larger than V_m then the STATCOM feeds reactive power into the system, and if V_o is smaller, it absorbs reactive power.

Referring to Fig. 22-37*a*, the amplitudes of the sending end, midpoint, and receiving end voltages are assumed to be equal for simplicity ($V_s = V_m = V_r = V$). The STATCOM compensation at the midpoint effectively segments the transmission line into two independent parts, each with an effective line reactance of *X*/2. Neglecting losses, the real power flow is the same in both parts, and can be derived as given in Eq. (22-57).

$$P = \frac{V^2}{(X/2)} \sin\left(\frac{\delta}{2}\right) \tag{22-57}$$

where δ is the angle between the sending and receiving end voltages. Figure 22-37*b* shows the variation of real power flow with phase angle, as determined by Eq. 22-57. The curve corresponding to no compensation is also shown for comparison. It can be clearly seen that the STATCOM significantly improves the transient stability margin, that is, for a given fault clearing time, STATCOM allows a much higher real power to be transmitted. In the case of power system oscillations, such as inter area oscillations, the shunt compensation is varied dynamically to provide damping.

Static Synchronous Series Compensator. SSSC is a series connected device that injects a synchronous line frequency voltage, normally in quadrature with the line current. The SSSC controls power flow by controlling the line voltage amplitude, phase angle, and effective line impedance.

Unified Power Flow Controller. The UPFC is a versatile FACTS device that combines the functions of a STATCOM and an SSSC, and extends their capability to inject shunt current or series voltage that involve real power flow as well [79]. With UPFC, the real and reactive power can be controlled independently. Unified power flow controller is capable of controlling all the power system parameters such as voltage magnitudes, phase angles, and effective line impedance simultaneously and therefore meet multiple control objectives. Figure 22-38*a* shows the schematic diagram of the UPFC. It consists of two voltage source converters with separate controllers but sharing a common dc link with dc storage capacitors.

In present installations of UPFC, most of the control functions are performed by the series converter by injecting a voltage V_{inj} whose phase is independent of the line current and can vary practically from 0° to 360°. The magnitude of the injected voltage can also be varied continuously within the rating of the series converter. The main function of the shunt converter is to provide the real power exchanged by the series converter with the system. It may be noted that the real power exchanged by the series converter is ultimately derived from the transmission line, but the reactive power is

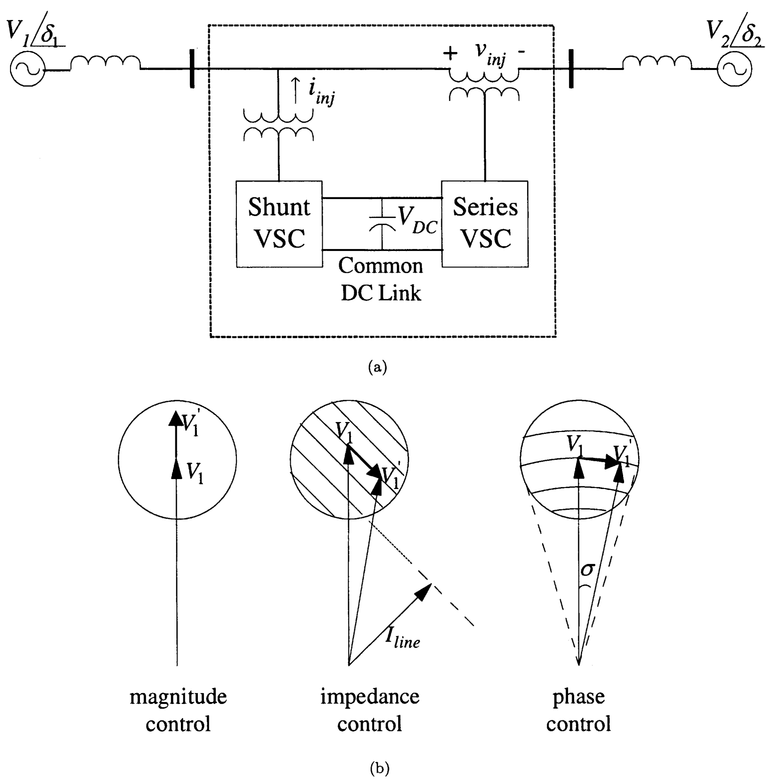


FIGURE 22-38 Unified power flow controller: (*a*) schematic, (*b*) control capabilities.

absorbed or supplied locally by the series converter and does not need to come from the transmission system. The shunt converter can be operated at unity power factor or can be controlled to provide additional functions beyond supporting real power needs of the series converter.

Figure 22-38*b* highlights the capabilities of the series converter of the UPFC, namely, control of voltage magnitudes, phase angles, and impedance [75]. With the injected voltage in phase or antiphase with the line voltage the UPFC provides voltage regulation or magnitude control. For a given voltage rating of the series converter, the in-phase addition provides maximum voltage magnitude control. For impedance control, the magnitude of the injected voltage is proportional to the line current and the phase is in quadrature (leading or lagging). In phase angle control, the magnitude and angle of the injected voltage are controlled such that the sending end voltage has the required phase angle without any change in the magnitude. These three features can be combined to achieve multifunction power flow control.

22.10.3 Custom Power

The digital age loads, for example, processing plants in semiconductor industry and data centers, require clean and uninterrupted power. These loads are highly intolerant to (even momentary) power quality problems such as voltage sags or interruptions, harmonics in line voltage, phase unbalance, and flicker in supply voltage. Power electronic systems that mitigate power quality problems in utility distribution systems (1 to 38 kV) are defined as custom power devices [80–82].

Similar to FACTS, the custom power devices can be connected in shunt or series with the distribution line or a combination of both. The major custom power devices are the dynamic voltage restorer (DVR), distribution static compensator (DSTATCOM) and the unified power quality controller (UPQC). The DVR is a series connected device that injects a controlled voltage to compensate for voltage sags and other momentary disturbances. The DSTATCOM is a shunt-connected device injecting controlled currents at the point of common coupling to compensate for power quality problems in the load current. Unified power quality controller combines the features of DVR and DSTATCOM.

Dynamic Voltage Restorer. Short duration voltage sags are the predominant power quality events, with estimated revenue lost per event of more than $1 million for pharmaceutical industries. Power acceptability curves that quantify voltage disturbances in terms of magnitude of these sags (and swells) and duration of the disturbance have been developed [83]. The most popular of the power acceptability curves is the Computer Business Equipment Manufacturers Association (CBEMA) curve shown in Fig. 22-39. This was developed by the CBEMA, now the Information Technology Industry Council (ITIC). The semiconductor industry has its own standard called SEMI F47, developed by the Semiconductor Equipment and Materials Institute (SEMI).

Dynamic voltage restorers, which are among the most installed custom power devices, protect sensitive equipment against short term voltage disturbances [84]. Figure 22-40 shows the schematic of a DVR. As seen, DVR is a voltage source converter based series connected device that injects a line frequency voltage of appropriate magnitude and phase such that the voltage across the sensitive loads is always well regulated, and any disturbances in the input voltage is not propagated to the load. The voltage source converter is implemented using IGBT switches, which operate at frequencies in the range of tens of kilohertz. They have fairly high control bandwidth and can respond to voltage disturbances in a small fraction of the line frequency cycle.

When the DVR is not connected to an active dc source and cannot handle real power in steady state, the injected voltage is constrained to be in phase quadrature with the load current. With this mode of control, the magnitude of sags that a DVR can correct becomes a function of the load power factor, and at higher power factor (close to unity) only a smaller voltage disturbance can be corrected. Since most of the sag events are of short duration, many of the installed DVRs rely on the large energy storage capacitor to supply real power for a short duration, and not constrain the injected voltage to be in phase quadrature. Several other installations use a separate rectifier to supply real power to the dc link from the distribution system [85].

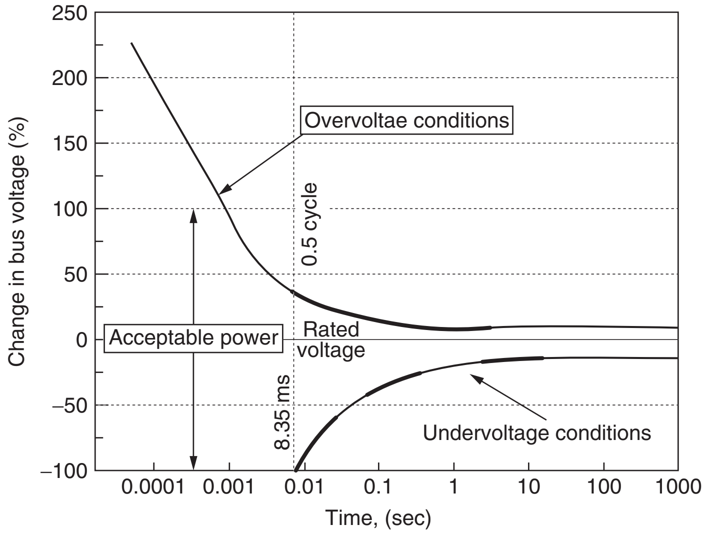


FIGURE 22-39 The CBEMA curve.

Distribution Static Compensator. The Distribution STATCOM has similar structure as that of the STATCOM used in transmission systems, and injects controlled currents. However, the main objectives of DSTATCOM are quite different. The load currents in distribution system can be unbalanced and contain reactive and harmonic components. Standards such as IEEE 519 and IEC 61000 place limits on maximum permissible harmonic currents for various types of equipment and voltage levels [47, 48]. The DSTATCOM with closed loop control injects correction currents such that the compensated load draws balanced, fundamental, unity power factor current.

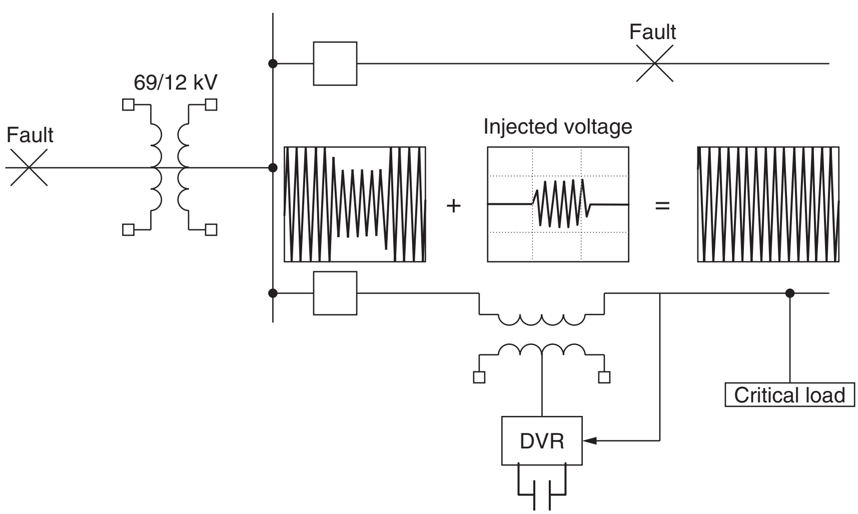


FIGURE 22-40 Application of a DVR.

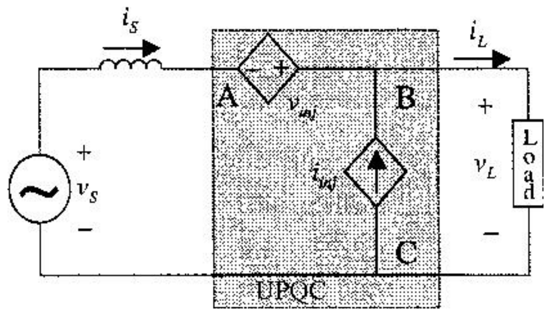


FIGURE 22-41 Unified power quality conditioner.

Unified Power Quality Conditioner. The UPQC combines the features of a DVR and DSTATCOM and can inject current in shunt and voltage in series simultaneously. Figure 22-41 shows the schematic of a UPQC. It has the same structure as that of UPFC used in transmission systems, consisting of two voltage source converters sharing a common dc link. One of the converters is connected in series with the distribution line injecting controlled voltages and the other converter is connected in shunt and injects controlled currents. Therefore, the UPQC can simultaneously correct for unbalances and distortion in line voltage as well as load currents.

Solid-state switches used to connect critical loads to multiple feeders or to break short circuit currents, hence improving power quality, are also considered as custom power devices. These are referred to as network reconfiguring devices and include solid-state current limiter (SSCL), solid-state breaker (SSB) and solid-state transfer switch (SSTS) [80]. These are much faster than the conventional mechanical switches and hence significantly enhance the reliability of the distribution system.

22.10.4 Distribution Generation Interface

Distributed generation (DG) is an important energy option for the twenty-first century and a key element of the restructuring of electric grid. Distribution generation, using renewable sources such as solar, wind, and tidal energy, and hydrogen (with photovoltaics for hydrogen generation) represents one of the most promising paths to sustainable energy.

Fuel cells, photovoltaics (PV), wind energy, and microturbines are among the most promising distributed generation technologies at present. Most of the distributed energy resources (DER) require a power electronic converter to interface with the power system network. Fuel cells and PV require dc-ac conversion, microturbines require high frequency to line frequency conversion, and generators used to capture wind energy are controlled through a rotor side power converter. Further, these DER may also require power electronics controlled energy storage.

Figure 22-42 shows a schematic diagram of the converter that can be used to interconnect photovoltaics with the utility. The input voltage from the solar cell array, typically 52 to 90 V, is converted to a higher magnitude, well regulated and isolated dc voltage through a high-frequency dc-dc converter. The dc link voltage is then converted to the required 60 Hz ac voltage by using a PWM

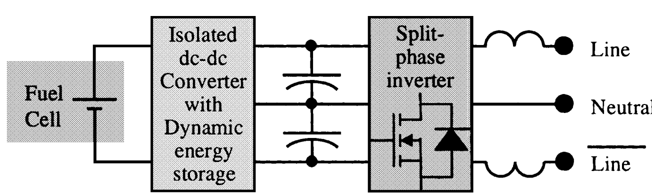


FIGURE 22-42 Fuel cell connection to the grid.

voltage source converter. An important feature of a PV interface is the maximum power point tracking circuitry designed to derive the maximum possible energy from solar radiation by suitably adjusting the current drawn from the solar panel. Fuel cells can also be interfaced to the grid by using a similar configuration as shown in Fig. 22-42. Since fuel cells have very little overcurrent or short-circuit rating, an energy storage capacitor is installed at the input side of the converter.

Wind energy is widely considered to be the fastest growing alternate energy source. The cost of wind energy in large wind farms located in good wind sites, at about 5 cents/kWh, is now competitive with the conventional utility generation, and is expected to reduce further. In 2004, the total installed capacity of wind energy worldwide was above 46,000 MW, with the United States accounting for 7000 MW [86].

Large wind farms consist of several multi megawatt wind turbines that are interconnected with the utility grid through medium voltage collector network. Doubly fed induction generators (DFIG) with a wound rotor and an ac-dc-ac PWM converter, as shown in Fig. 22-43, is becoming the most widely used technology for wind generation. The main advantage of the DFIG based generation is that it allows extraction of maximum energy from the wind at varying wind speeds, with the power converter rated only for about 15% of the total power. The stator winding is connected directly to the utility grid while the rotor is supplied with controlled, variable frequency currents by the PWM converter. By appropriate control of the rotor currents, the machine can generate power from subsynchronous to super synchronous speeds. Another advantage of DFIG is that the grid side converter can generate or absorb reactive power.

Apart from serving as an environment-friendly energy source, the DG systems are expected to provide various other benefits. For example, the concept of microgrids is gaining prominence [87]. Microgrid is a cluster of DER with power converters, energy storage and loads, which can be controlled together and present to the grid as a single entity. With suitably designed power converters, and with coordinated control, the microgrids can enhance stability, provide relief to transmission congestion, and provide reactive power support. Another promising approach is to use distributed micro sources for combined heat and power (CHP) [88]. Due to the various ancillary functions expected, inverter based DG are becoming more widespread compared to traditional reciprocating engines.

The interconnect standards for DG are just being evolved, like the IEEE 1547 standard for interconnecting distributed resources with electric power systems [89]. IEEE 1547 specifies the requirements for the DG to disconnect from the grid under deviations in voltage magnitude and frequency or under grid outages. Much work is still needed to realize a universal interconnect technology (UIT) and to understand the effect of large penetration of DG on fault currents, protection and dynamic interactions with the power system.

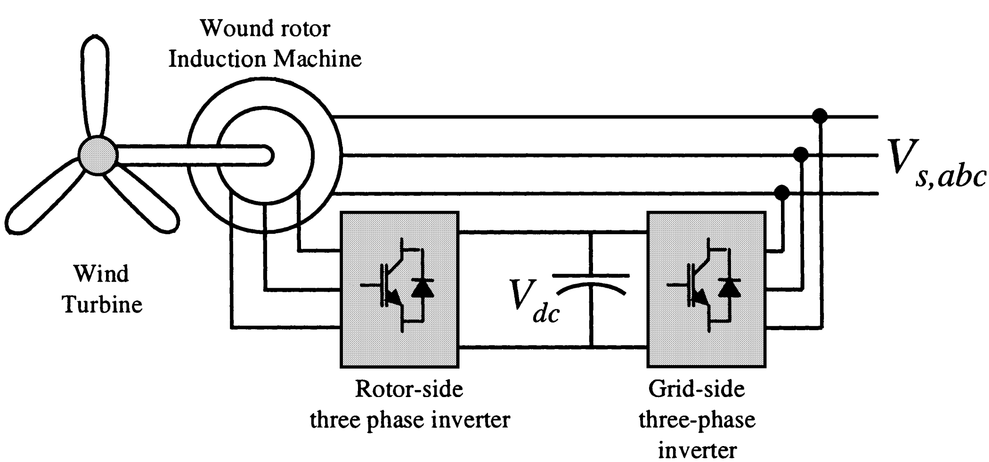


FIGURE 22-43 Grid connection of a doubly fed induction generator.

22.11 COMPONENTS OF POWER ELECTRONIC CONVERTERS

This section describes the individual components that constitute a power converter. These include power semiconductor devices and passive components (inductors, transformers, capacitors). A detailed description of the structure and physics of power semiconductor devices is beyond the scope of this discussion. The interested reader is referred to standard books on this subject [4, 90, 91]. Details of magnetics, material properties, and capacitors are also omitted. Unlike semiconductor devices and capacitors, very few magnetics are available as standard products from manufacturers, and therefore require custom design. Magnetics design is covered in significant detail in Ref. [92]. For details of capacitors, application notes and datasheets supplied by manufacturers have to be relied on.

22.11.1 Power Semiconductor Devices

Power electronic circuits require high-power semiconductor switches and diodes. An ideal switch should have the following characteristics: full control over switch state (on/off), very low voltage drop during on state, infinite impedance during off-state, and instantaneous transition between states. Diodes should have very low voltage drop during conduction, infinite impedance in off-state, and instantaneous transition. Practical devices have nonideal characteristics, and different devices capitalizing on one advantage while sacrificing some other have been developed. Figure 22-44 shows the circuit symbols of common power semiconductor devices. These are listed below with their voltage, current, and switching limitations.

Diodes: Line frequency, fast recovery, ultra-fast recovery, and schottky—in increasing order of switching speed, and decreasing order of reverse voltage rating.

MOSFETs (metal oxide semiconductor field effect transistor): Good for low voltage ~100s of volts, high switching frequency (> 100 kHz).

IGBTs (insulated gate bipolar transistors): Good from a few hundred Volts to about 6 kV, currents upto 1.2 kA, and switching frequency upto 30 kHz.

Thyristors or SCRs (silicon-controlled rectifiers): Good for very high voltage and current (~kV and kA), and low power moderate performance applications.

GTOs (gate turn-off thyristors): Good for very high voltage (~kV), high current applications (~kA), with switching frequency upto a few kilohertz.

Miscellaneous: IGCT (integrated gate commutated thyristor), MCT (MOS-controlled thyristor), BJTs (bipolar junction transistors), triacs, diacs.

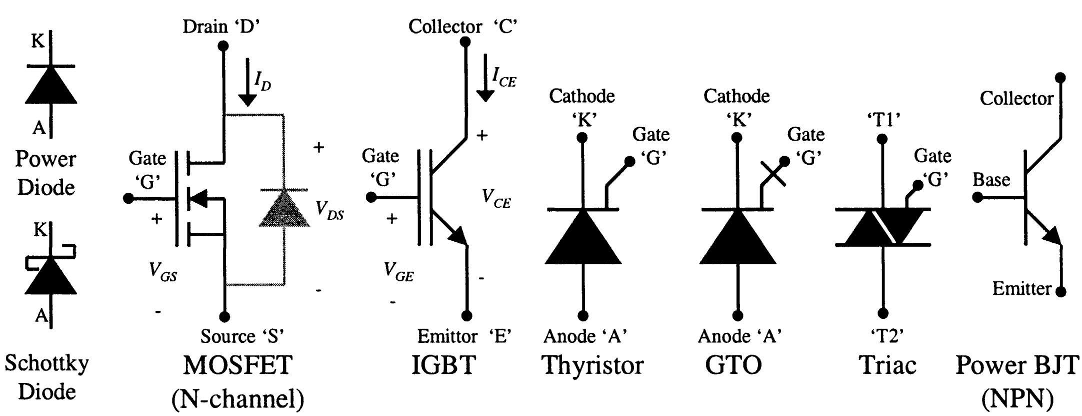


FIGURE 22-44 Circuit symbols of common power semiconductor devices.

Diodes. There are essentially two types of power semiconductor diodes: PN junction, and metal-semiconductor junction (schottky). PN junction power diodes have an additional N^- (lightly doped with N type impurities) drift region, so the overall structure is PN^-N. The depletion layer of the PN junction extends in the N^- region when the diode is reverse biased, and its length determines the maximum reverse voltage the diode can block. Intrinsically, the N^- region has a high resistance. However, when the diode is forward biased there is injection of excess carriers in this region resulting in a low effective resistance. This phenomena is commonly called *conductivity modulation*. The on-state voltage drop across the diode consists of the PN junction drop and the resistive drop in the drift region; typically it ranges from 0.6 to 1 V. There is a small delay in going from off-state to on-state due to the time required for the carriers to build up. During turn-off, the excess carriers in the drift region have to be removed. Thus, for a short time, the diode conducts in the reverse direction with a high voltage across it. This phenomena, known as reverse recovery, leads to significant power loss and becomes one of the limiting factors in high-frequency circuits. PN junction type power diodes are classified as:

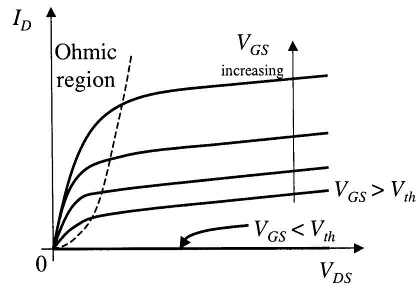
- line frequency rectifiers: for rectification of 50/60 Hz utility input.
- fast and ultra-fast recovery diodes: for high frequency rectification. These have recovery time ranging from a couple of μs to fractions of μs.

Schottky diodes are based on metal-semiconductor junctions. These junctions have a lower junction potential leading to a lower forward voltage drop. Silicon-based schottky power diodes have forward voltage drop ranging from 0.3 to 0.6 V, and can withstand reverse voltages up to 200 V. As opposed to PN junction diodes, schottkys are majority carrier devices, so they do not have any reverse recovery. However, compared to PN junction power diodes, they have significantly higher capacitance. This capacitance, in combination with circuit inductances, can lead to significant oscillations when the diode goes from on-state to off-state. Silicon-based schottky diodes are suited for very high frequency, low voltage, and high current rectification. Recently, SiC-based schottky diodes have been developed and are now commercially available with rating up to 1200 V [93, 94]. Even though SiC diodes have a much higher forward voltage drop (2 to 3 V), the absence of reverse recovery makes them suitable for high voltage high frequency rectification.

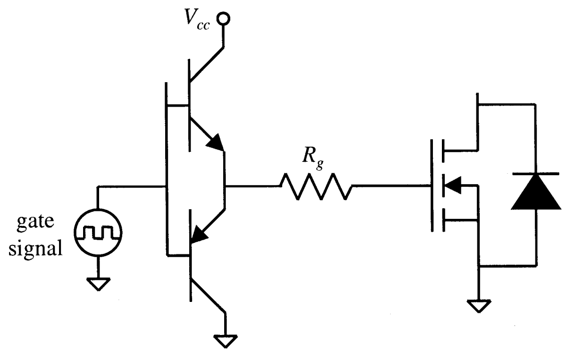
MOSFETs. Unlike signal level MOS devices that are fabricated laterally, power MOSFETs have a vertically diffused structure. For an N-channel MOSFET the doping is of the form $N^+PN^-N^+$. The drain is the N^+ terminal next to the N^- region, while the source is the N^+ region next to the P region. A positive voltage on the isolated gate terminal produces an electron channel in the P-region allowing current to flow from drain to source (or vice versa). Further, the P-type body is shorted to the source terminal resulting in an intrinsic diode with anode at the source and cathode at the drain. Although this diode is not very good in performance, it is useful for most power electronic circuits. Both P and N channel power MOSFETs are available, but N-channel MOSFETs are more prevalent due to their lower on-state resistance. Our discussion will therefore be restricted to N-channel MOSFETs.

The steady-state V-I characterisitcs for an N-channel MOSFET are shown in Fig. 22-45*a*. For $V_{GS} < V_{th}$, the MOSFET acts as an open circuit from drain to source; V_{th}, the threshold voltage, is in the range of 2 to 4 V. For $V_{GS} > V_{th}$, the MOSFET follows the characteristics shown in Fig. 22-45*a*. In amplifier circuits, MOSFETs are operated in their active region, where the drain current I_D is almost independent of the drain to source voltage V_{DS}. Power MOSFETs are operated in the ohmic region where I_D is proportional to V_{DS}, and the MOSFET behaves like a resistance. The effective on-resistance, designated R_{DS}, depends on V_{GS}, I_D, and the junction temperature T_j. For MOSFETs, R_{DS} increases with increasing temperature, a property useful in paralleling of devices to obtain higher current carrying capacity. The maximum value of V_{GS} is usually ± 20 V; for logic level power MOSFETs, it is limited to ± 10 V. For most MOSFETs, increasing V_{GS} beyond 10 V does not have significant effect on R_{DS}. MOSFETs are rated for (1) maximum drain to source breakdown voltage (BV_{DSS}) (2) maximum continuous average current for a specified temperature (e.g., I_{D25} at 25°C), and (3) a safe operating area (SOA) in terms of V_{DS}, I_D, and time duration for which I_D flows. In very high current applications, thermal performance of the device package can also impose an additional constraint.

MOSFET gate drive circuits usually run on a nominal V_{cc} = 15 or 12 V, and switch V_{GS} between V_{cc} and 0 V. For very high current applications and improved noise sensitivity, a negative V_{GS} may



(a)



(b)

FIGURE 22-45 MOSFETs: (*a*) V-I characteristics, and (*b*) simple gate drive circuit.

be applied during the off-time. Since the gate is insulated from the source there is no dc current flow from gate to source. However, depending on their ratings power MOSFETs have a significant input capacitance C_{iss}. Thus, to increase V_{GS} from 0 to 15 V in a very short time (20 to 100 ns) a significant current (order of 1 A) is required. Figure 22-45*b* shows a basic MOSFET gate drive circuit. Several gate drive ICs that can source and sink current upto a few amperes, while providing several other auxiliary functions, are commercially available [9, 12]. For *high-side switches*, switches whose source voltage changes with the switch state, as is the case for the MOSFET in a buck converter, an isolated gate drive is required. The isolation is provided by using either high-frequency transformers or high-speed optocouplers. There are several standard gate drive configurations, each with its own pros and cons. The best source of these circuits are application notes from device manufacturers (e.g., see Ref. [62]).

Power MOSFETs can be used up to a few 100 kHz and in some applications in the megahertz range. To increase operating frequencies further, *RF Power MOSFETs* have recently been introduced [95]. These can be operated in the 10 MHz range and are expected to reduce converter size and weight considerably.

Insulated Gate Bipolar Transistors. The structure of IGBTs is similar to MOSFETs. However, IGBTs have significantly higher voltage and current ratings compared to MOSFETs, and their on-state voltage drop is also lower. In addition, several IGBT chips are paralleled inside one package to form an IGBT module with a significantly higher current rating. Insulated gate bipolar transistors modules are commercially available in voltage rating upto 6.5 kV, and currents up to 1200 A. The gate drive requirements of an IGBT are similar to that of MOSFETs. It is turned on by applying a positive gate source voltage (typically 15 V), and turned-off by applying a smaller negative voltage (about −5 V). Compared to MOSFETs, IGBTs have much longer switching times. Insulated gate bipolar transistors are minority carrier devices, so their turn-off is characterized by a "tail current". Their switching frequencies are generally limited to 30 kHz, and the maximum switching frequency reduces with power level. If resonant power conversion techniques are used, the tail current during turn-off can be avoided and the switching frequency can be pushed higher.

Thyristor and Similar Devices. Thyristors, also called silicon-controlled rectifiers (SCRs), are high-power semiconductor devices that can block voltage of either polarity and conduct current in one direction only (from anode to cathode). They can be switched on by applying a current pulse to their gate terminal (with return path through cathode) when there is a positive voltage from anode to cathode. They can be switched off only by reducing the device current to zero. Thyristors are available in very high current and voltage ratings and have a very low conduction voltage drop. Thyristors are mostly used for ac to dc (or vice-versa) power conversion, where the device current is expected to reduce to zero. Normally, thyristors are switched at or close to the ac system frequency. They are ideally suited for utility applications such as HVDC and controlled reactors.

Gate turn-off thyristors are very high power devices, with their low-end ratings overlapping with the high-end rating of IGBTs. Unlike a thyristor, they can be turned on and off using the gate terminal, although the gate drive is more complex compared to a MOSFET or an IGBT. The switching times for GTOs are of the order of 10 μs so their maximum switching frequency is in the kilohertz range. They are used exclusively in very high power applications like some motor drives, FACTs devices, and active filters. Gate turn-off are commercially available in ratings upto 6 kV and 6 kA. Enhancements to GTO have led to the development of IGCT (integrated gate commutated thyristor) [96] or GCT (gate commutated turn-off thyristor) [97], ETO (emitter turn-off thyristor), and MTO (MOS turn-off thyristor) [91]. At present there is a strong research effort in development of high power devices with high switching speeds. The impetus behind these are high power pulse applications for defence, and the increasing utility applications of power converters.

A "back-to-back connection" of two thyristors (anode of one to the cathode of other) can conduct current in either direction and block voltages of either polarity. Triacs realize the functionality in a single semiconductor device. However, voltage and current ratings of triacs are very low compared to that of thyristors.

Power *bipolar junction transistors* (BJTs) have been almost completely replaced by MOSFETs and IGBTs, due to ease of control and higher switching frequency. However, they are still used in some applications like linear power supplies.

22.11.2 Magnetic Components

In power electronic converters three types of magnetic components are used: single winding inductors (for filtering current and aiding in resonant transitions in some circuits), multi-winding coupled inductors (to provide filtering and isolation), and transformers (for isolation and stepping up/down voltage). Unlike semiconductor devices, these have to be custom designed using available cores, wires, etc. The primary consideration for their design are size/weight and power loss. Design procedures using the common *area product method* is presented here. Another approach is the *core geometry method*, which can be found in Refs. [5, 92]. It is assumed that the reader is familiar with basics of electromagnetism and magnetic circuits. An E-E type core will be used for illustration, but the method is applicable to any core shape. The E-E type core construction, along with relevant definitions of its geometry, is shown in Fig. 22-46. The complete core is formed with two E cores with possibly an air gap between them. The coils are wound over a plastic bobbin placed on the outside of the center leg of the E sections.

Due to high-frequency magnetic fields, significant eddy currents are induced in the magnetic core and windings, and due to high-frequency electric field there can be significant capacitive currents between windings. Eddy currents lead to significant losses and necessitate the use of high-resistivity magnetic core materials and thinner wires (which may be paralleled for high current capacity) for the coils. Loss due to magnetic hysteresis also increases with increase in frequency.

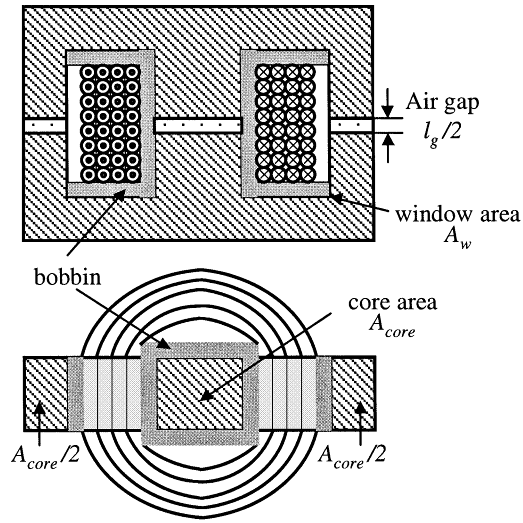


FIGURE 22-46 Vertical and horizontal sections of an E-E core inductor.

Magnetic Core Materials. Inductors use one of three different kinds of core materials, depending on the currents they are supposed to carry. The three core material types are silicon steel for low-frequency filtering, powdered iron for high-frequency filtering, and ferrites for carrying high-frequency currents. Transformers in power converters always carry high-frequency currents and therefore use ferrite cores.

For low-frequency filtering, where the inductor primarily carries a dc component with a 120 Hz ripple (as in line frequency rectifiers), standard silicon steel laminations can be used to form the core. Silicon steel has high saturation flux density (~1.8 Tesla) and high relative permeability (μ_r~ 40,000). This results in a small core size without incurring significant core loss. However, the relative permeability reduces with increase in flux density, so the effective inductance value changes with the dc component of the current.

For inductors that are supposed to carry currents consisting of a large dc component and a small high-frequency ac component (e.g., the inductor used at the output of a dc-dc converter), powdered iron or MolyPermalloy Powder (MPP) cores are used. These cores have high resistivity leading to lower eddy current losses, high saturation flux density (upto 1.4 Tesla) leading to lower core size, and distributed air gap leading to a low relative permeability. As with silicon steel, the relative permeability reduces with increasing flux density. The cores are available in toroidal and E shapes. Toroids are difficult to wind but can be used to make very good quality inductors.

For transformers and inductors that have to carry significant high-frequency current, the powdered iron material has unacceptable loss due to eddy currents and hysteresis. Ferrite materials, with very high resistivity, lower weight, but low saturation flux density (~ 0.4 Tesla) are suitable for these. Common ferrite materials are 3F3 from ferroxcube and PC44 from TDK. Cores made from these materials are available in a wide variety of shapes and sizes: toroids, pot core, E, PQ, RM, etc.

Inductor Design. For inductors, the design requires specified values of the inductance (L), and the peak and rms current ($\hat{I}$ and I_{rms}) it has to carry. The relations between peak current and core area (A_{core}), rms current and window area (A_w), and the expression for the *area product*, $A_p = A_{core} \cdot A_w$, are as follows:

$$\lambda = L\hat{I} = N\Phi = NB_{max} \cdot A_{core}$$
$$\Rightarrow A_{core} = \frac{L\hat{I}}{NB_{max}} \tag{22-58}$$

$$I_{rms} = J \cdot A_{cond} = J \cdot \frac{A_w \cdot k_w}{N}$$
$$\Rightarrow A_w = \frac{NI_{rms}}{Jk_w} \tag{22-59}$$

$$A_p = A_w \cdot A_{core} = \frac{L\,\hat{I}I_{rms}}{JB_{max}k_w} \tag{22-60}$$

Here, λ is the flux linkage, N is the number of turns of the winding, Φ is the magnetic flux corresponding to peak current $\hat{I}$, B_{max} is the maximum flux density (may be different from the saturation flux density), J is the current density in the winding, A_{cond} is the cross-sectional area of one conductor in the winding, and k_w called the winding factor is an empirical factor used to indicate the fraction of window area utilized by the copper of the windings. In the expression for A_p, the numerator consists of specified quantities, while the denominator has quantities which are chosen for design. B_{max} may be chosen close to the saturation flux density depending on allowable core loss; J is chosen in the range of 2.5 to 8 A/mm^2, depending on allowable conduction loss, mechanism for heat removal, and the allowable temperature rise; k_w is in the range of 0.3 to 0.8 depending on space taken by bobbin and insulation, and the method of winding. Once the area product has been calculated, a core with the desired value of A_p can be chosen from manufacturers' catalogs. The number of turns and the air gap required to achieve the desired inductance value are then calculated as

$$N = \frac{L\hat{I}}{B_{max} A_{core}} \tag{22-61}$$

$$L = N^2/\Re = N^2\mu_o A_{core}/l_g \tag{22-62}$$

$$\Rightarrow l_g = N^2\mu_o A_{core}/L \tag{22-63}$$

where $\Re$ is the core reluctance. The above equation assumes that $\mu_r/l_m \gg 1/l_g$, where l_m is the magnetic path length of the core. The conductor size is calculated as $A_{cond} = k_w \cdot A_w/N$. If the current is expected to have significant high frequency component, the radius of the conductors should not exceed the *skin depth* of copper at the expected frequency. Several conductors may then be paralleled to achieve the required value of A_{cond}. Finally, P_{cond}, conduction loss in the windings, and P_{core}, loss in the core, should be calculated. P_{cond} requires an estimate of the winding resistance, while P_{core} is computed using manufacturer supplied empirical loss curves. Core loss in an inductor depends on the operating flux density B (related to the average value of the current), and change in flux density during each switching cycle ΔB (related to the expected current ripple). It is generally accepted that $P_{cond} = P_{core}$ indicates a good design. However, power loss in the core can be dissipated more easily compared to windings (especially if there are several layers). Thus, it may be desirable to have $P_{cond} < P_{core}$. This affects the choice of B_{max} and J values used above.

Powdered iron cores have distributed air gaps and the above procedure is not directly applicable to them. For selection of these cores, the manufacturers suggest a number of cores based on a metric similar to the product $L\hat{I}I_{rms}$. The final core selection can then be made based on allowable losses (P_{cond} relating to J, and P_{core} relating to B_{max}), and temperature rise. The cores have a specified A_L value, the inductance obtained from the core in nH/Turns^2. The number of turns required to obtain the inductance value the can be computed using the specified A_L value. Finally, the inductance variation with dc current should be checked to make sure that the required value is obtained at the maximum operating current.

Transformer Design. For a transformer, the design specifications are the turns ratio, rms current in all the windings, and the maximum volt-second product ([$V\,sec$], the maximum product of voltage and time for which the voltage is applied) expected across the primary winding. The window area can be related to the rms current, the core area to the maximum volt-second product, and the resulting area product calculated as follows

$$\begin{aligned} k_w \cdot A_w &= N_p \cdot A_{cond,p} + \sum_m (N_{sm} \cdot A_{cond,s}) \\ \Rightarrow A_w &= \frac{N_p}{Jk_w} \cdot \left(I_{p,rms} + \sum_m [(N_{sm}/N_p) \cdot I_{sm,rms}]\right) \end{aligned} \tag{22-64}$$

$$\begin{aligned} \Delta\lambda &= [V\,sec] = N_p \Delta\Phi = N_p \Delta B_{max} \cdot A_{core} \\ \Rightarrow A_{core} &= \frac{[V\,sec]}{N_p \Delta B_{max}} \end{aligned} \tag{22-65}$$

$$A_p = A_w \cdot A_{core} = \frac{[V\,sec] \cdot (I_{p,rms} + \sum_m [(N_{sm}/N_p) \cdot I_{sm,rms}])}{J\Delta B_{max} k_w} \tag{22-66}$$

The subscripts "*p*" and "*s*" indicate primary and secondary, respectively, while the index *m* refers to the secondary number. In the above equations, ΔB_{max} is the maximum *change* in flux density; this equals the maximum value of B_{max} for single quadrant operation (as in a forward converter), and twice B_{max} when flux direction reverses (as in full-bridge dc-dc converters). Core loss has more than quadratic dependence on ΔB_{max}. For transformers, the value of B_{max} is therefore chosen much lower than the saturation flux density (about 0.1 to 0.25 T). After a core is chosen by utilizing the computed A_p, the number of primary turns are calculated as

$$N_p = \frac{[V\,sec]}{B_{max} A_{core}} \qquad (22\text{-}67)$$

In high-frequency transformers, the magnetizing current can be a significant part of the primary current and should be accounted for in the above calculations. Conductor selection for transformers is similar to that for inductors. For very high frequency and high current applications use of insulated copper foils and *Litz wire* are alternatives to use of multiple parallel conductors.

It should be noted that the procedures described above are rudimentary and several other factors have to be accounted for in a real design. These include ambient temperature, temperature rise of the core (which affects the saturation flux density and the core loss significantly), *hot-spot* temperature, winding loss due to *fringing fields* near the air gap, and winding loss due to *proximity effect*. In transformers, interleaving primary and secondary windings helps in reducing the proximity loss and leakage inductance at the expense of increased interwinding capacitance. To eliminate the winding process, planar magnetics that use tracks on multilayer PCBs to form the windings have been developed. Planar magnetics also have the benefit of low profile. To reduce the number of discrete magnetic components, there has been significant research in integrated magnetics—realizing the functionality of several magnetic components using a single magnetic component.

22.11.3 Capacitors

There are several types of capacitors each with different advantages and limitations. For power electronic circuits the choice criteria are: capacitance value, voltage rating, effective series resistance (ESR) and effective series inductance (ESL), and the ripple and ac current capacity at different frequencies. All these can be found as specifications in manufacturers datasheets. ESR is either specified as a number at different frequencies (typically 120 Hz and 20/100 kHz), or in terms of the *dissipation factor*, $\tan\delta = ESR/(2 \times f \times C)$. If ESL is specified, it may be either as a number or in terms of the series resonance frequency, $f_r = 1/(2\pi\sqrt{ESL \times C})$. The main capacitor types used in power electronic circuits are listed in the Table 22-3. The list is by no means exhaustive and there are several other capacitor types suited to different applications, for example, tantalum, mica, and high power film capacitors.

22.11.4 Snubber Circuits

As described in Sec. 22.2.4, there is significant energy loss in power semiconductor devices during turn-on and turn-off. In addition to the phenomena described there, reverse recovery in PN junction based diodes and capacitance of schottky diodes, output capacitance of MOSFETs and IGBTs, and leakage inductance of transformers also contribute to power loss during switching. Snubber circuits are used to remove the transient high voltage and high current stress from semiconductor devices. They can be classified into two basic types—*turn-on* and *turn-off* snubbers, depending on the transition where they are active. With a turn-off snubber, an alternate path is provided for the current during turn-off so that the switch ceases to carry current as soon as it is disabled. This does not necessarily imply that the switching loss is eliminated. In low-power applications, the energy may be dissipated in an external resistor and in fact the energy dissipated may be higher. For high-power applications, the energy may be recovered using energy recovery snubbers, resulting in significantly improved

TABLE 22-3 Capacitors Types Commonly Used in Power Electronic Circuits

Capacitor Type	Characteristics	Application
Electrolytic	Polarized (unipolar voltage only) High density, High ESR and ESL Low reliability Voltage ratings up to 500 V Capacitance up to 100s of mF Rated for ripple current capacity	DC input and output filters, very short time energy storage
Ceramic and Multilayer Ceramic	Very low ESL and ESR Low capacitance values (up to ~100 μF) Maximum capacitance value reduces with increasing voltage rating	Paralleled for low voltage output filtering, "bypass" for gate drives, high voltage capacitors (up to 1 kV) used in snubber circuits
(Metallized) Polyester film	Low ESR and ESL Low capacitance values (up to a few μF) High voltage ratings (few kV) Bigger than ceramic and electrolytic	In input and output dc filters: to suppress switching transients, ac filters, snubber circuits
Metallized Polypropylene film	Very low ESR and ESL Low capacitance values (up to a few μF) Very high rms current and high voltage ratings	In resonant converters: for carrying high frequency ac current

efficiency at the expense of increased component count. Another use of snubber circuits is to limit transient overvoltages caused due to parasitics.

22.11.5 Heat Sinks

The main power loss in power electronic circuits occurs in power semiconductor devices, windings, and cores of the magnetic components, capacitors subjected to high-ripple current, and auxiliary circuits like gate drives. The resulting heat has to be removed from the components and eventually transferred to the atmosphere/ambient surroundings due to the following reasons. Power semiconductors have a maximum junction temperature rating beyond which they fail, and on-state resistance of devices like MOSFETs increases with junction temperature leading to reduced efficiency. Electrolytic capacitors can fail when heated beyond their ratings and their expected lifetime reduces with increase in temperature. Increased temperatures in magnetics can lead to higher power loss, poor magnetic core characteristics, insulation breakdown, and shorter lifetimes.

Most semiconductor devices are cooled by mounting their package on a heat sink. A heat sink is essentially a piece of metal designed to dissipate heat by convection and radiation (by maximizing the heat sink surface area). The heat transfer from the device to the heat sink is via conduction. If the package of the device is not electrically isolated from its electrical terminals, an interface material (like Silpads) with good thermal conductivity and high dielectric breakdown voltage has to be used. For magnetics, resins may be used to conduct heat from the windings and the core to the metal heat sink. The heat sink itself may be cooled naturally, by forced convection (using fans to blow air on the heat sink surface), or liquid cooled (by circulating a liquid through the mass of the heat sink, and cooling the liquid by a radiator). Simple steady analysis and design of heat sinks is carried out using specified *thermal resistances*, ratio of temperature difference to heat transferred across the material, of all the materials in the heat transfer path. For pulsed applications, a transient thermal model, usually consisting of first-order lags between two consecutive interfaces, is used.

REFERENCES

1. J. G. Kassakian, J. M. Miller, and N. Traub, "Automotive electronics power up," *IEEE Spectrum*, vol. 37, no. 5, pp. 34–36, May 2000.
2. R. Middlebrook and S. Ćuk, *Advances in Switched-mode Power Conversion.* CA: TESLAco, 1981, vol. 1.
3. H. Khalil, *Nonlinear Systems,* 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.
4. N. Mohan and W. P. Robbins, *Power Electronics: Converters, Applications, and Design,* 3rd ed. Hoboken, NJ: John Wiley & Sons, Incorporated, 2003.
5. R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*, 2nd ed. New York, NY: Springer-Verlag, 2001.
6. A. I. Pressman, *Switching Power Supply Design*, 2nd ed. New York: McGraw-Hill, 1997.
7. K. Billings, *Switchmode Power Supply Handbook*, 2nd ed. New York: McGraw-Hill, 1999.
8. R. P. Severns and G. E. Bloom, *Modern DC-to-DC Switchmode Power Converter Circuits,* 1st ed. New York: Van Nostrand Reinhold Company, 1985.
9. Texas instruments company website. [Online]. Available: http://www.ti.com/
10. National semiconductors company website. [Online]. Available: http://www.national.com/
11. Maxim company website. [Online]. Available: http://www.maxim-ic.com/
12. Linear technologies company website. [Online]. Available: http://www.lti.com/
13. R. Ayyanar and N. Mohan, "Novel soft-switching dc-dc converter with full zvs-range and reduced filter requirement part i: Regulated-output applications," *IEEE Transactions on Power Electronics*, vol. 16, no. 2, pp. 184–192, Mar. 2001.
14. —, "Novel soft-switching dc-dc converter with full zvs-range and reduced filter requirement part ii: Constant-input, variable-output applications," *IEEE Transactions on Power Electronics*, vol. 16, no. 2, pp. 193–200, Mar. 2001.
15. On semiconductor company website. [Online]. Available: http://www.onsemi.com/
16. Power integrations company website. [Online]. Available: http://www.powerint.com/
17. Texas instruments power management webpage. [Online]. Available: http://www.ti.com
18. S. R. Sanders and G. C. Verghese, "Lyapunov-based control for switched power converters," *IEEE Transactions on Power Electronics*, vol. 7, no. 1, pp. 17–24, Jan. 1992.
19. V. Vorpérian, "Simplified analysis of pwm converters using model of a pwm switch part i: continuous conduction mode," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 3, pp. 490–496, May 1990.
20. —, "Simplified analysis of pwm converters using model of a pwm switch part ii: discontinuous conduction mode," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 3, pp. 497–505, May 1990.
21. N. Mohan, *First Course on Power Electronics and Drives*. Minneapolis, MN: MNPERE, 2003.
22. SPICE. University of Berkeley. [Online]. Available: http://www-cad.eecs.berkeley.edu/Software/software.html
23. K. Ogata, *Modern Control Engineering*, 4th ed. NJ: Prentice Hall, 2001.
24. H. D. Veneble, "The k facor: a new mathematical tool for stability analysis and synthesis," in *Proc. of the International Solid-State Power Electronics Conference (POWERCON 10)*, San Diego, CA, Mar. 1983, pp. 1–12.
25. *ORCAD PSpice A/D User's Guide*, Cadence Design Systems, 2003.
26. G. C. Verghese, *Nonlinear Phenomena in Power Electronics: Bifurcations, Chaos, Control, and Applications.* Hoboken, NJ: John Wiley & Sons, 2001.
27. M. P. Kazmierkowski, R. Krishnan, and F. Blaabjerg, *Control in Power Electronics: Selected Problems*. Elsevier Science & Technology Books, 2002.
28. B. K. Bose, *Power Electronics and Variable Frequency Drives: Technology and Applications.* NJ: IEEE Press, 1996.
29. —, *Modern Power Electronics and AC Drives.* NJ: Prentice Hall, 2001.
30. N. Mohan, *Advanced Electric Drives.* Minneapolis, MN: MNPERE, 2001.
31. R. C. Panaitescu and N. Mohan, "A simple space-vector pwm algorithm for vsi-fed ac motor drives," in *Proc. IEEE Applied Power Electronics Conference and Exposition (APEC'02)*, vol. 1, Dallas, TX, Mar. 2002, pp. 72–75.

32. H. Krishnamurthy, G. Narayanan, R. Ayyanar, and V. T. Ranganathan, "Design of space vector-based hybrid pwm techniques for reduced current ripple," in *Proc. IEEE Applied Power Electronics Conference and Exposition (APEC'03)*, vol. 1, Miami Beach, FL, Feb. 2003, pp. 583–588.
33. F. Blashke, "The principle of field orientation as applied to the new transvector closed loop control system for rotating field machines," *Siemens Review*, vol. 34, no. 5, pp. 217–220, 1972.
34. M. Depenbrock, "Direct self-control (dsc) of inverter-fed induction machine," *IEEE Transactions on Power Electronics*, vol. 3, no. 4, pp. 420–429, 1988.
35. I. Takahashi and T. Noguchi, "A new quick-response and high efficiency control strategy of an induction motor," *IEEE Transactions on Industry Applications*, vol. 22, no. 5, pp. 820–827, 1986.
36. P. Vas, *Sensorless Vector and Direct Torque Control*. New York: Oxford University Press, 1998.
37. D. W. Novotony and T. A. Lipo, *Vector Control and Dynamics of AC Drives*. Oxford: Claredon Press, 1996.
38. K. Corzine, *The Power Electronics Handbook, Chapter 6: Multilevel Converters*, T. L. Skvarenina, Ed. New York, NY: CRC Press, 2002.
39. J. Rodriguez, J. Lai, and F. Peng, "Multilevel inverters: a survey of topologies, controls, and applications," *IEEE Transaction on Industrial Electronics*, vol. 40, no. 4, pp. 724–738, Aug. 2002.
40. B. P. McGrath, D. G. Holmes, and T. Lipo, "Optimized space vector switching sequences for multilevel inverters," *IEEE Transactions on Power Electronics*, vol. 18, no. 6, pp. 1293–1301, Nov. 2003.
41. A. Nabae, I. Tkahashi, and H. Akagi, "A new neutral-point clamped pwm inverter," *IEEE Transactions on Industry Applications*, vol. IA-17, no. 5, pp. 518–523, Sept.-Oct. 1981.
42. L. M. Tolbert, F. Z. Peng, and T. G. Habetler, "Multilevel converters for large electric drives," *IEEE Transactions on Industry Applications*, vol. 35, no. 1, pp. 36–44, Jan.-Feb. 1999.
43. G. K. Dubey, S. R. Doradla, A. Joshi, and R. M. K. Sinha, *Thyristorised Power Controllers*. New York: John Wiley & Sons, 1986.
44. K. R. Padiyar, *HVDC Power Transmission Systems*. New Delhi, India: Wiley Eastern, 1993.
45. P. W. Wheeler, J. Rodrigueza, J. C. Clare, L. Empringham, and A. Weinstein, "Matrix converters: A technology review," *IEEE Transactions on Industry Electronics*, vol. 49, no. 2, pp. 276–278, Apr. 2002.
46. A. Alesina and M. G. B. Venturini, "Analysis and design of optimum-amplitude nine-switch direct ac-ac converters," *IEEE Transactions on Power Electronics*, vol. 4, no. 1, pp. 101–112, Jan. 1989.
47. *IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems*, IEEE Std. 519-1992, 1992.
48. *Electromagnetic Compatibility Part 3: Limits*, International Electrotechnical Commission (IEC) Std. 1000-3, 1995.
49. R. Ayyanar and A. K. Jain, *The Power Electronics Handbook, Section 17.5: Unity Power Factor Rectification*, T. L. Skvarenina, Ed. New York, NY: CRC Press, 2002.
50. M. J. Nave, *Power Line Filter Design for Switched-Mode Power Supplies*. Van Nostrand Reinhold, 1991.
51. Federal communications commission. [Online]. Available: http://www.fcc.gov
52. Verband deutscher elektrotechniker (association of german electrical engineers). [Online]. Available: http://www.vde.com/vde_en
53. Artesyn technologies company website. [Online]. Available: http://www.artesyn.com/powergroup
54. N. Mohan, *Electric Drives: An Integrative Approach*. Minneapolis, MN: MNPERE, 2003.
55. T. J. Miller, Ed., *Brushless Permanent-Magnet and Reluctance Motor Drives*. Oxford: Oxford University Press, 1989.
56. T. J. E. Miller, *Switched Reluctance Motors and their Control*. Oxford: Magna Physics Publishing and Clarendon Press, 1993.
57. T. J. E. Miller, Ed., *Electronic Control of Switched Reluctance Machines*. Oxford: Newnes, 2001.
58. Powerex, inc. company website. [Online]. Available: http//www.pwrx.com
59. Semikron company website. [Online]. Available: http://www.semikron.com
60. Mitshubishi electric semiconductor website. [Online]. Available: http://www.mitsubishichips.com/Global/index.html
61. Eupec company website. [Online]. Available: http://www.eupec.com
62. International recitifier company website. [Online]. Available: http://www.irf.com

63. W. F. Ray and R. M. Davis, "Inverter drive for doubly salient reluctance motor: its fundamental behaviour, linear analysis and cost implications," *IEE Journal on Electric Power Applications*, vol. 2, no. 6, pp. 185–193, Dec. 1979.
64. A. K. Jain, "Two phase modeling, experimental characterization, and power converter with fast demagnetization for switched reluctance motor drives," Ph.D. dissertation, University of Minnesota, Minneapolis, MN, 2003.
65. M. Barnes and C. Pollock, "Power electronic converters for switched reluctance drives," *IEEE Transactions on Power Electronics*, vol. 13, no. 6, pp. 1100–1111, Nov. 1998.
66. S. Vukosavic and V. R. Stefanovic, "Srm inverter topologies: a comparative evaluation," *IEEE Transactions on Industry Applications*, vol. 27, no. 6, pp. 1034–1047, Nov.-Dec. 1991.
67. B. Masoud, "Charging ahead with li-ion and li-pol batteries," *Power Electronics Technology*, vol. 28, no. 7, pp. 26–29, July 2002.
68. Building technologies program: Lighting research and development website. U. S. Department of Energy. [Online]. Available: http://www.eere.energy.gov/buildings/tech/lighting/
69. R. L. Steigerwald, "A comparison of half-bridge resonant converter topologies," *IEEE Transactions on Power Electronics*, vol. 3, no. 2, pp. 174–182, Apr. 1988.
70. P. Greenland and W. Burns, "Powering next-generation solid-state lighting," *Power Electronics Technology*, vol. 30, no. 5, pp. 34–39, May 2004.
71. J. G. Kassakian, "Automotive electrical systems-the power electronics market of the future," in *Proceedings of the Annual IEEE Applied Power Electronics Conference and Exposition (APEC'00)*, vol. 1, New Orleans, LA, Feb. 2000, pp. 3–9.
72. J. Cooper and A. Agarwal, "Sic power-switching devices-the second electronics revolution?" *Proceedings of the IEEE*, vol. 90, no. 6, pp. 956–968, June 2002.
73. A. Elasser and T. P. Chow, "Silicon carbide benefits and advantages for power electronic circuits and systems," *Proceedings of the IEEE*, vol. 90, no. 6, pp. 969–986, June 2002.
74. N. G. Hingorani, "Flexible ac transmission," *IEEE Spectrum*, vol. 30, no. 4, pp. 40–45, Apr. 1993.
75. N. G. Hingorani and L. Gyugyi, *Understanding Facts: Concepts and Technology of Flexible AC Transmission Systems*. New York: John Wiley & Sons, 1999.
76. P. Kundur, *Power System Stability and Control*. New York: McGraw-Hill, 1994.
77. L. Gyugyi, "Application characteristics of converter-based facts controllers," in *Proc. International Conference on Power System Technology (PowerCon'00)*, vol. 1, Perth, WA, Australia, Dec. 2000, pp. 391–396.
78. K. Sen, "Statcom-static synchronous compensator: theory, modeling, and applications," in *Proceedings of the IEEE Power Engineering Society Winter Meeting*, vol. 2, New York, NY, Jan.-Feb. 1999, pp. 1177–1183.
79. L. Gyugyi, C. Schauder, S. Williams, T. Rietman, D. Torgerson, and A. Edris, "The unified power flow controller: a new approach to power transmission control," *IEEE Transactions on Power Delivery*, vol. 10, no. 2, pp. 1085–1097, Apr. 1995.
80. A. Ghosh and G. Ledwich, Eds., *Power quality enhancement using custom power*. Boston, MA: Kluwer Academic Publishers, 2002.
81. N. G. Hingorani, "Introducing custom power," *IEEE Spectrum*, vol. 32, no. 6, pp. 41–48, June 1995.
82. *IEEE Standard on Custom Power, in Draft Stage*, IEEE Std. P1409.
83. J. Kyei, R. Ayyanar, G. Heydt, R. Thallam, and J. Blevins, "The design of power acceptability curves," *IEEE Transactions on Power Delivery*, vol. 17, no. 3, pp. 828–833, July 2002.
84. J. Nielsen, M. Newman, H. Nielsen, and F. Blaabjerg, "Control and testing of a dynamic voltage restorer (dvr) at medium voltage level," *IEEE Transactions on Power Electronics*, vol. 19, no. 3, pp. 806–813, May 2004.
85. N. Woodley, "Field experience with dynamic voltage restorer (dvrtmmv) systems," in *Proc. IEEE Power Engineering Society Winter Meeting*, vol. 4, New York, NY, Jan.-Feb. 2000, pp. 2864–2871.
86. J. Flowers. (2005, Jan.) Wind powering America update. WPA. [Online]. Available: http://www.eere.energy.gov/windandhydro/windpoweringamerica/pdfs/wpa/wp%a_update.pdf
87. R. Lasseter, "Microgrids," in *Proceedings of the IEEE Power Engineering Society Winter Meeting*, vol. 1, New York, NY, Jan. 2002, pp. 305–308.
88. T. M. Kaarsberg and J. M. Roop, "Combined heat and power: how much carbon and energy can manufacturers save?" *Aerospace and Electronic Systems Magazine*, vol. 14, no. 1, pp. 7–12, Jan. 1999.

89. *IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems*, IEEE Std. 1547, 2003.
90. J. B. Baliga, *Power Semiconductor Devices*. New York: PWS Publishing Company, 1995.
91. T. L. Skvarenina, Ed., *The Power Electronics Handbook*, *Part I: Power Electronic Devices*, New York: CRC Press, 2002.
92. W. T. McLyman, *Transformer and Inductor Design Handbook*, 3rd ed. New York: Marcel Dekker, 2002.
93. Infineon technologies company website. [Online]. Available: http://www.infineon.com
94. Cree semiconductors company website. [Online]. Available: http://www.cree.com
95. Ixys rf company website. [Online]. Available: http://www.ixysrf.com
96. Abb Power semiconductors website. ABB. [Online]. Available: http://www.abb.com
97. Mitsubishi electric power semiconductors website. Mitushi Electric. [Online]. Available: http://www.mitsubishichips.com/Global/products/power/index.html

SECTION 23
POWER QUALITY AND RELIABILITY

Surya Santoso
Senior Member IEEE, Assistant Professor, Electrical and Computer Engineering, University of Texas at Austin

Mark F. McGranaghan
Senior Member IEEE, Associate Vice President, EPRI Solutions, Inc., Knoxville, TN.

Roger C. Dugan
Fellow IEEE, Senior Consulting Engineer, EPRI Solutions, Inc., Knoxville, TN.

CONTENTS

23.1 PERSPECTIVE ON POWER QUALITY

23.1.1 Introduction

Power quality is about compatibility between the quality of the voltage supplied from the electric power system and the proper operation of end-use equipment. Power quality is also about economics—finding the optimum level of investment in the power system and the end-use equipment to achieve the compatibility. There are two categories of power quality that need to be considered—steady-state (or continuous) power quality and disturbances. Steady-state power quality characteristics include voltage regulation, harmonic distortion, unbalance, and flicker. We can define *compatibility levels* for these characteristics and then the challenge is to maintain performance within these compatibility levels and make sure that equipment can operate with these levels. Power quality disturbances (outages, momentary interruptions, voltage sags, and transients) are much more of a challenge. It is impossible to completely prevent disturbances that may cause equipment disruptions. Therefore, we have to find the best balance between investments to prevent disturbances and investments in equipment and facility protection.

On the technology side, future power quality research will focus on advanced technologies that can be applied at all levels of the system to improve compatibility (both supply-side technologies and end-user technologies) and on the procedures to find the optimum places to make these investments from a system perspective. The result will be guidance regarding expected levels of performance for different types of supply systems that will result in optimum economics if customers also make the associated investments to assure that the required equipment performance. Recommendations from the economic analysis will also require regulatory structures to support the implementation of optimum system designs and solutions. Therefore, the research results must be coordinated with development of regulations and market structures for future power systems.

23.2 CATEGORIES AND CHARACTERISTICS OF POWER QUALITY DISTURBANCE PHENOMENA

23.2.1 General

Power quality is a generic term applied to a wide variety of electromagnetic phenomena on the power system. The duration of these phenomena ranges from a few nanoseconds (e.g., lightning strokes) to a few minutes (e.g., feeder voltage regulations) to steady-state disturbances (harmonic distortions and voltage fluctuations). Due to the extensive variety of the phenomena, many power quality terms have sometimes been applied incorrectly and cause confusion among end users, vendors, and service providers in dealing with power quality concerns. For example, a term *power surge* has been used to describe some kind of power disturbances. However, it is ambiguous and in fact has no technical meaning since power surge does not refer to a surge in power. This term has been used to refer to overvoltage transients in voltage. Power is related to the product of voltage and current. Normally, *voltage* is the quantity causing the observed disturbance and the resulting power will not necessarily be directly proportional to the voltage. The solution will generally be to correct or limit the voltage as opposed to addressing the power. Therefore, the use of ambiguous and nonstandard terms is discouraged.

23.2.2 General Classes of Power Quality Disturbances

The Institute of Electrical and Electronics Engineers Standards Coordinating Committee 22 (IEEE SCC22) has led the main effort in the United States to coordinate power quality standards. It has the responsibilities across several societies of the IEEE, principally the Industry Applications Society and the Power Engineering Society. It coordinates with international efforts through liaisons with the IEC and CIGRE (International Conference on Large High-Voltage Electric Systems). The IEC classifies electromagnetic phenomena into the groups shown in Table 23-1[1].

The U.S. power industry efforts to develop recommended practices for monitoring electric power quality have added a few terms to the IEC terminology [2]. *Sag* is used as a synonym to the IEC term *dip*. The category *short duration variations* is used to refer to voltage dips and short interruptions. The term *swell* is introduced as an inverse to sag (dip). The category *long duration variation* has been added to deal with American National Standards Institute (ANSI) C84.1 limits. The category *noise* has been added to deal with broadband conducted phenomena. The category *waveform distortion* is used as a container category for the IEC *harmonics, interharmonics*, and *dc in ac networks* phenomena as well as an additional phenomenon from IEEE Std. 519-1992 (Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems), called *notching*. Table 23-2 shows the categorization of electromagnetic phenomena used for the power quality community.

23.2.3 Transient—General

The term *transient* has long been used in the analysis of power system variations to denote an event that is undesirable and momentary in nature. Other definitions in common use are broad in scope and simply state that a transient is "that part of the change in a variable that disappears during transition from one steady-state operating condition to another" [8]. Another word in common usage that is often considered synonymous with transient is *surge*. This term should be avoided unless it is qualified with appropriate explanation. In general, transients can be classified into two categories, impulsive and oscillatory. These terms reflect the waveshape of a current or voltage transient.

TABLE 23-1 Principal Phenomena Causing Electromagnetic Disturbances as Classified by the IEC

Conducted low-frequency phenomena
Harmonics, interharmonics
Signal systems (power line carrier)
Voltage fluctuations (flicker)
Voltage dips and interruptions
Voltage imbalance (unbalance)
Power-frequency variations
Induced low-frequency voltages
DC in ac networks
Radiated low-frequency phenomena
Magnetic fields
Electric fields
Conducted high-frequency phenomena
Induced continuous wave (CW) voltages or currents
Unidirectional transients
Oscillatory transients
Radiated high-frequency phenomena
Magnetic fields
Electric fields
Electromagnetic fields
Continuous waves
Transients
Electrostatic discharge phenomena (ESD)
Nuclear electromagnetic pulse (NEMP)

Impulsive Transient. An *impulsive transient* is a sudden, nonpower frequency change in the steady-state condition of voltage, current, or both, that is unidirectional in polarity (primarily either positive or negative). They are normally characterized by their rise and decay times which can also be revealed by their spectral content. For example, a 1.2 × 50 μs 2000-V impulsive transient nominally rises from zero to its peak value of 2000 V in 1.2 μs, and then decays to half its peak value in 50 μs. The most common cause of impulsive transient is lightning. Figure 23-1 illustrates a typical current impulsive transient caused by lightning.

Oscillatory Transient. An *oscillatory transient* is a sudden, nonpower frequency change in the steady-state condition of voltage, current, or both, that includes both positive and negative polarity values. It consists of a voltage or current whose instantaneous value changes polarity rapidly. It is described by its spectral content (predominate frequency), duration, and magnitude. The spectral content subclasses defined in Table 23-2 are high, medium, and low frequency. The frequency ranges for these classifications are chosen to coincide with common types of power system oscillatory transient phenomena. High- and medium-frequency oscillatory transients are transients with a primary frequency component greater than 500 kHz with a typical duration measured in microseconds, and between 5 and 500 kHz with duration measured in the tens of microseconds, respectively. Figure 23-2 illustrates a medium frequency oscillatory transient event due to back-to-back capacitor energization.

23.2.4 Short-Duration Voltage Variations

Short-duration voltage variations are caused by fault conditions, the energization of large loads that require high starting currents, or intermittent loose connections in power wiring. Depending on the fault location and the system conditions, the fault can cause either temporary voltage drops (sags), or voltage rises (swells), or a complete loss of voltage (interruptions). The fault condition can be close to or remote from the point of interest. In either case, the impact on the voltage during the actual fault condition is of short duration variation until protective devices operate to clear the fault.

TABLE 23-2 Categories and Characteristics of Power System Electromagnetic Phenomena

Categories	Typical Spectral Content	Typical Duration	Typical Voltage Magnitude
1.0 Transients			
1.1 Impulsive			
1.1.1 Nanosecond	5 ns rise	< 50 ns	
1.1.2 Microsecond	1 μs rise	50 ns–1 ms	
1.1.3 Millisecond	0.1 ms rise	> 1 ms	
1.2 Oscillatory			
1.2.1 Low frequency	< 5 kHz	0.3–50 ms	0–4 pu*
1.2.2 Medium frequency	5–500 kHz	20 μs	0–8 pu
1.2.3 High frequency	0.5–5 MHz	5 μs	0–4 pu
2.0 Short duration variations			
2.1 Instantaneous			
2.1.1 Interruption		0.5–30 cycle	< 0.1 pu
2.1.2 Sag (dip)		0.5–30 cycle	0.1–0.9 pu
2.1.3 Swell		0.5–30 cycle	1.1–1.8 pu
2.2 Momentary			
2.2.1 Interruption		30 cycle–3 s	< 0.1 pu
2.2.2 Sag (dip)		30 cycle–3 s	0.1–0.9 pu
2.2.3 Swell		30 cycle–3 s	1.1–1.4 pu
2.3 Temporary			
2.3.1 Interruption		3 s–1 min	< 0.1 pu
2.3.2 Sag (dip)		3 s–1 min	0.1–0.9 pu
2.3.3 Swell		3 s–1 min	1.1–1.2 pu
3.0 Long duration variations			
3.1 Interruption, sustained		> 1 min	0.0 pu
3.2 Undervoltages		> 1 min	0.8–0.9 pu
3.3 Overvoltages		> 1 min	1.1–1.2 pu
4.0 Voltage unbalance		steady state	0.5–2%
5.0 Waveform distortion			
5.1 DC offset		steady state	0–0.1%
5.2 Harmonics	0–100 Hz	steady state	0–20%
5.3 Interharmonics	0–6 kHz	steady state	0–2%
5.4 Notching		steady state	
5.5 Noise	broadband	steady state	0–1%
6.0 Voltage fluctuations	< 25 Hz	intermittent	0.1–7% 0.2–2 Pst
7.0 Power frequency variations		< 10 s	

*pu = per unit.

This category encompasses the IEC category of voltage dips and short interruptions. Each type of variation can be designated as instantaneous, momentary, or temporary, depending on its duration as defined in Table 23-2.

Interruption. An interruption occurs when the supply voltage or load current decreases to less than 0.1 pu for a period of time not exceeding 1 min. Interruptions can be the result of power system faults, equipment failures, and control malfunctions. The interruptions are measured by their duration since

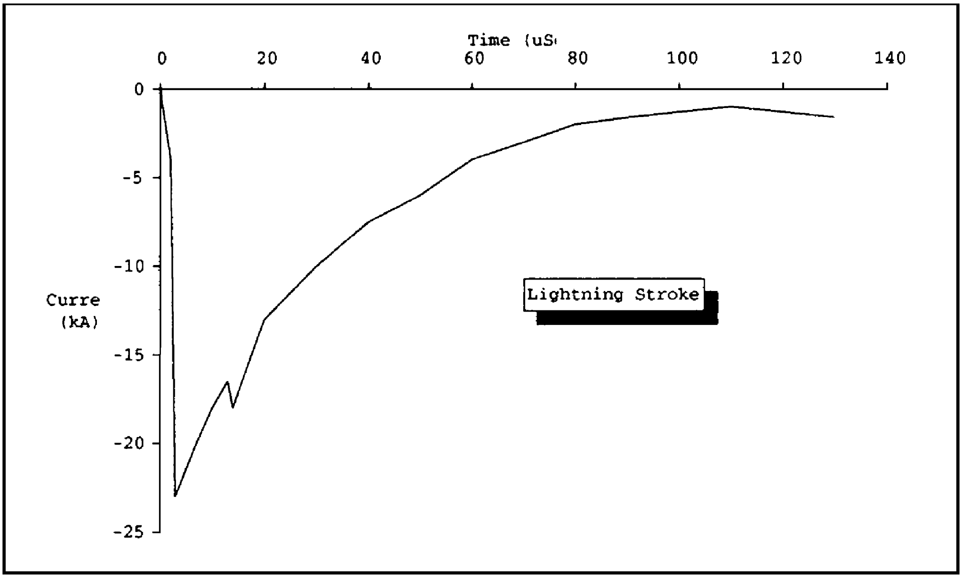


FIGURE 23-1 Lightning stroke current impulsive transient.

the voltage magnitude is always less than 10% of nominal. The duration of an interruption due to a fault on the utility system is determined by the operating time of utility protective devices. Instantaneous reclosing generally will limit the interruption caused by a nonpermanent fault to less than 30 cycles. Delayed reclosing of the protective device may cause an instantaneous, momentary, or temporary interruption. The duration of an interruption can be irregular due to equipment malfunction or loose connections. Some interruptions may be preceded by a voltage sag when the

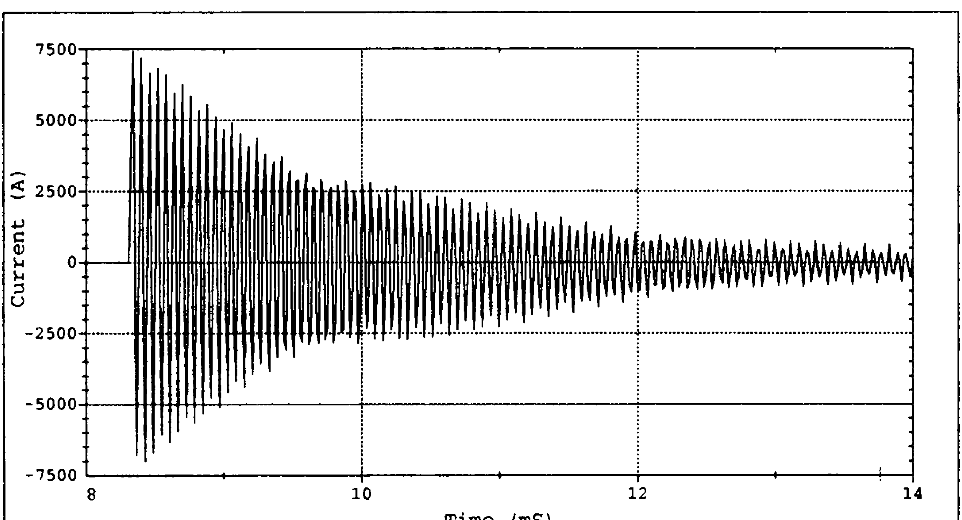


FIGURE 23-2 Oscillatory transient current caused by back-to-back capacitor switching.

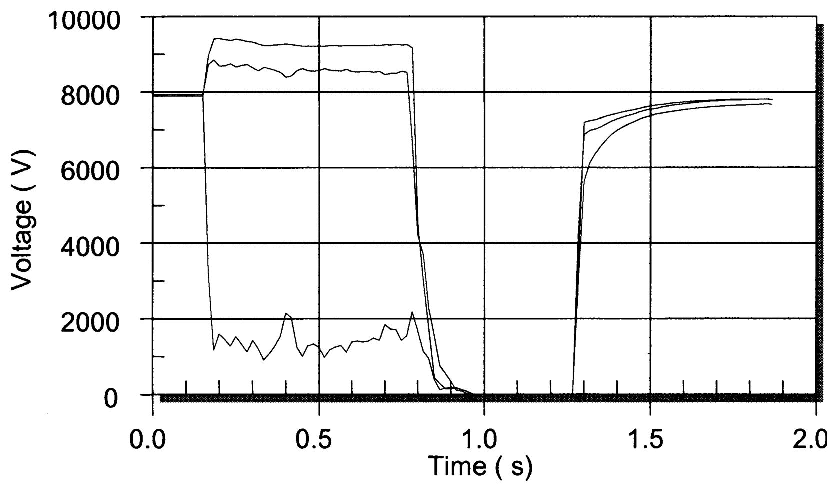


FIGURE 23-3 An instantaneous interruption due to a SLG fault and subsequent recloser operation.

interruptions are due to clearing faults on the source system. The voltage sag occurs between the time a fault initiates and the protective device operates. Figure 23-3 shows a plot of the rms voltages for all three phases for such an interruption. The voltage on the faulted phase initially sags to 15% to 25% for 0.6 s while the fault is arcing. A voltage swell occurs on the other two phases at the same time. The breaker then opens, clears the fault, and recloses successfully 0.4 s later. Utility distribution engineers frequently refer to this as an instantaneous reclose.

Voltage Sags. A sag is a decrease to between 0.1 and 0.9 pu in rms voltage or current at the power frequency for durations from 0.5 cycles to 1 min. The IEC definition for this phenomenon is *voltage dip*. The two terms are considered interchangeable, with sag being the preferred synonym in the U.S. power quality community. Figure 23-4 shows a typical voltage sag associated with a SLG fault on another feeder from the same substation. The voltage sags to 60% for about 5 cycles until the substation breaker is able to interrupt the fault current. Typical fault clearing times range from 3 to 30 cycles, depending on the fault current magnitude and the type of overcurrent protection.

Voltage Swells. A *swell* is defined as an increase to between 1.1 and 1.8 pu in rms voltage or current at the power frequency for durations from 0.5 cycle to 1 min. The term *momentary overvoltage* is used by many writers as a synonym for the term swell. As with sags, swells are usually associated with system fault conditions. One way that a swell can occur is from the temporary voltage rise on the unfaulted phases during a single line-to-ground (SLG) fault. An example is shown in Fig. 23-5. Swells can also be caused by switching off a large load or energizing a large capacitor bank.

23.2.5 Long-Duration Voltage Variations

Long-duration voltage variations encompass rms deviations at power frequencies for longer than 1 min. ANSI C84.1 specifies the steady-state voltage tolerances expected on a power system. A voltage variation is considered to be long duration when the ANSI limits are exceeded for greater than 1 min. Long-duration variations can be either overvoltages or undervoltages. Overvoltages and undervoltages generally are not the result of system faults, but are caused by load variations on the system and system switching operations. Such variations are typically displayed as plots of rms voltage versus time.

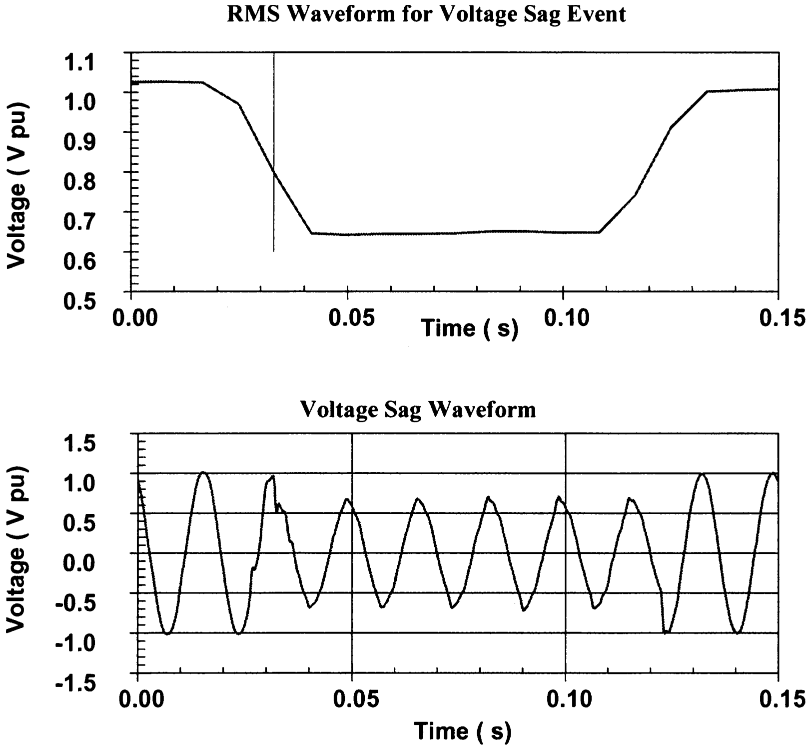


FIGURE 23-4 Voltage sag caused by a SLG fault.

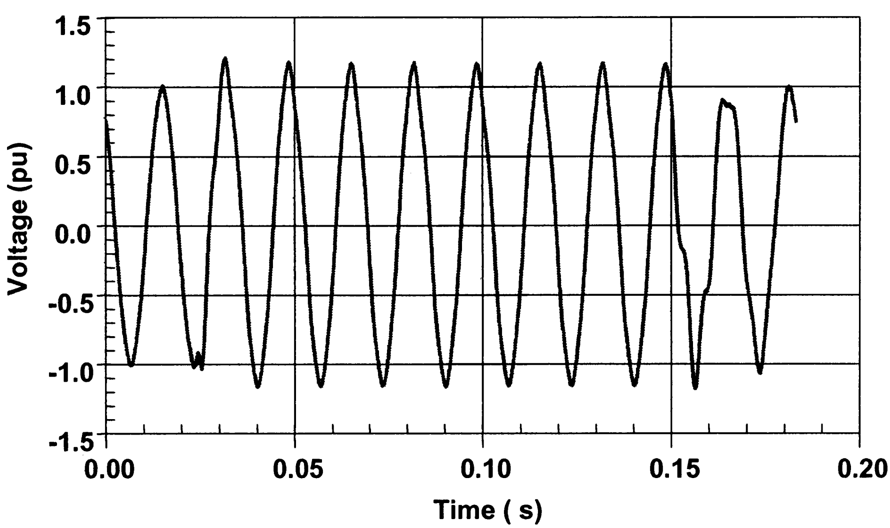


FIGURE 23-5 An 8-cycle voltage swell caused by a SLG fault.

Overvoltage. An *overvoltage* is an increase in the rms ac voltage greater than 110% at the power frequency for a duration longer than 1 min. They are usually the result of load switching (e.g., switching off a large load or energizing a capacitor bank). The overvoltages result because the system is either too weak for the desired voltage regulation or voltage controls are inadequate. Incorrect tap settings on transformers can also result in system overvoltages.

Undervoltage. An *undervoltage* is a decrease in the rms ac voltage to less than 90% at the power frequency for a duration longer than 1 min. They are the result of the events that are the reverse of the events that cause overvoltages. A load switching on or a capacitor bank switching off can cause an undervoltage until voltage regulation equipment on the system can bring the voltage back to within tolerances. Overloaded circuits can result in undervoltages also. The term *brownout* is often used to describe sustained periods of undervoltage initiated as a specific utility dispatch strategy to reduce power demand. Because there is no formal definition for brownout, and it is not as clear as the term undervoltage when trying to characterize a disturbance, the term brownout should be avoided.

23.2.6 Sustained Interruption

When the supply voltage has been zero for a period of time in excess of 1 min, the long duration voltage variation is considered a sustained interruption. Voltage interruptions longer than 1 min are often permanent and require human intervention to repair the system for restoration. The term sustained interruption refers to specific power system phenomena and, in general, has no relation to the usage of the term *outage*. Utilities use outage or interruption to describe phenomena of similar nature for reliability reporting purposes. However, this causes confusion for end users who think of an outage as any interruption of power that shuts down a process. This could be as little as one-half of a cycle. Outage, as defined in IEEE Std 100 [8], does not refer to a specific phenomenon, but rather to the state of a component in a system that has failed to function as expected. Also, use of the term *interruption* in the context of power quality monitoring has no relation to reliability or other continuity of service statistics. Thus, this term has been defined to be more specific regarding the absence of voltage for long periods.

23.2.7 Voltage Imbalance

Voltage imbalance (also called voltage unbalance) is sometimes defined as the maximum deviation from the average of the 3-phase voltages or currents, divided by the average of the 3-phase voltages or currents, expressed in percent. Unbalance is more rigorously defined in the standards [6, 8, 11, 12] using symmetrical components. The ratio of either the negative or zero sequence component to the positive sequence component can be used to specify the percent unbalance. The most recent standard [11] specifies that the negative sequence method be used. Figure 23-6 shows an example of these two ratios for a 1 week trend of imbalance on a residential feeder.

23.2.8 Waveform Distortion

Waveform distortion is defined as a steady-state deviation from an ideal sine wave of power frequency principally characterized by the spectral content of the deviation. There are five primary types of waveform distortion: dc offset, harmonics, interharmonics, notching, and noise.

DC Offset. The presence of a dc voltage or current in an ac power system is termed *dc offset*. This can occur as the result of a geomagnetic disturbance or asymmetry of electronic power converters. Incandescent lightbulb life extenders, for example, may consist of diodes that reduce the rms voltage supplied to the lightbulb by half-wave rectification. Direct current in alternating current networks can have a detrimental effect by biasing transformer cores so they saturate in normal operation. This

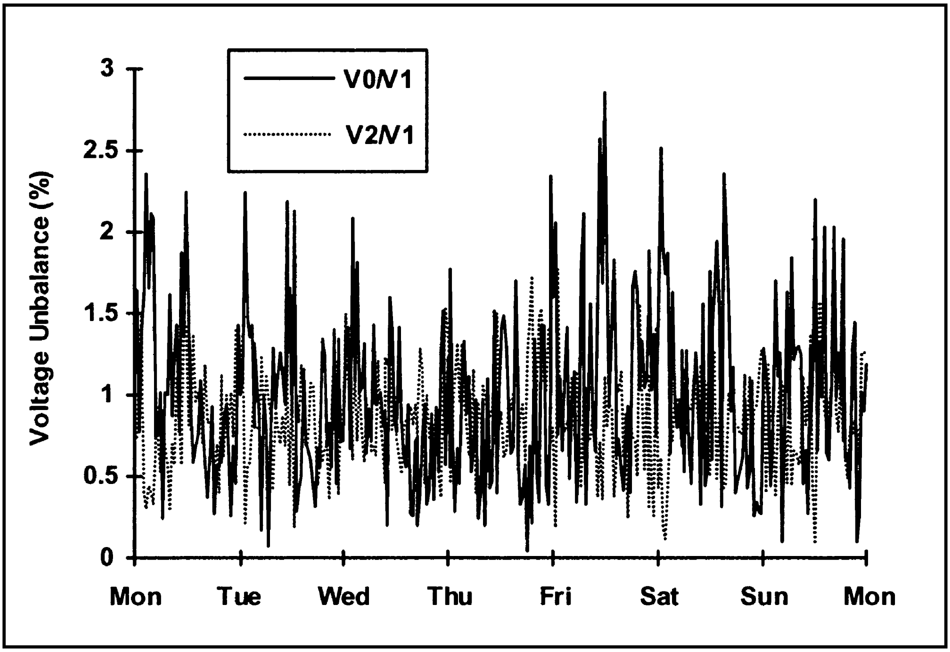


FIGURE 23-6 Voltage unbalance trend for a residential feeder.

causes additional heating and loss of transformer life. DC may also cause the electrolytic erosion of grounding electrodes and other connectors.

Harmonics and Interharmonics. *Harmonics* are sinusoidal voltages or currents having frequencies that are integer multiples of the frequency at which the supply system is designed to operate (termed the fundamental frequency; usually 50 or 60 Hz) [6]. Periodically distorted waveforms can be decomposed into a sum of the fundamental frequency and the harmonics. Harmonic distortion originates in the nonlinear characteristics of devices and loads on the power system. Figure 23-7 illustrates the waveform and harmonic spectrum for a typical adjustable speed drive input current.

Voltages or currents having frequency components that are not integer multiples of the frequency at which the supply system is designed to operate (e.g., 50 or 60 Hz) are called *interharmonics*. They can appear as discrete frequencies or as a wideband spectrum. Interharmonics can be found in networks of all voltage classes. The main sources of interharmonic waveform distortion are static frequency converters, cycloconverters, induction furnaces, and arcing devices. Power line carrier signals can also be considered as interharmonics.

Notching. *Notching* is a periodic voltage disturbance caused by the normal operation of power electronics devices when current is commutated from one phase to another. Since notching occurs continuously, it can be characterized through the harmonic spectrum of the affected voltage. However, it is generally treated as a special case. The frequency components associated with notching can be quite high and may not be readily characterized with measurement equipment normally used for harmonic analysis. Figure 23-8 shows an example of voltage notching from a 3-phase converter that produces continuous dc current. The notches occur when the current commutates from

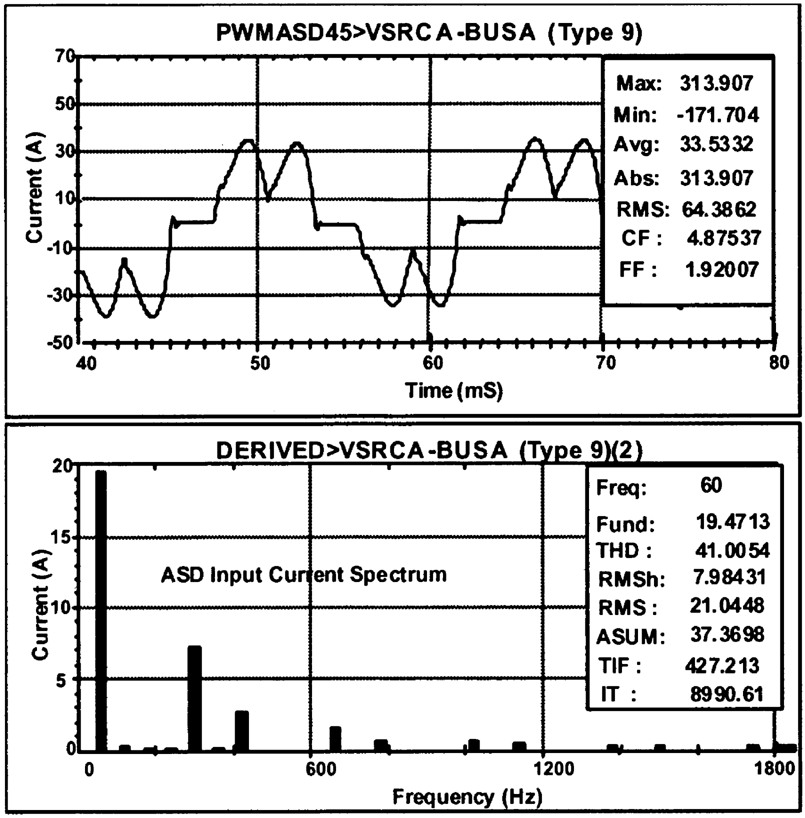


FIGURE 23-7 Current waveform and harmonic spectrum for an ASD input current.

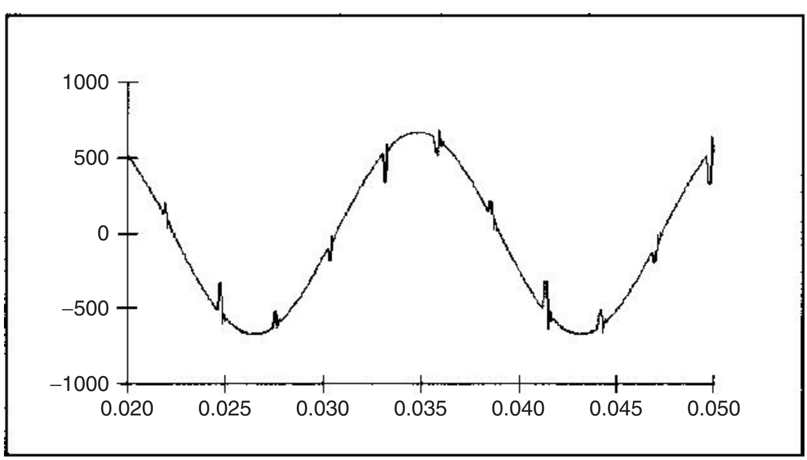


FIGURE 23-8 Example of voltage notching caused by a 3-phase converter.

one phase to another. During this period, there is a momentary short circuit between two phases pulling the voltage as close to zero as permitted by system impedances.

Noise. *Noise* is defined as unwanted electrical signals with broadband spectral content lower than 200 kHz superimposed upon the power system voltage or current in phase conductors, or found on neutral conductors or signal lines. Noise in power systems can be caused by power electronic devices, control circuits, arcing equipment, loads with solid-state rectifiers, and switching power supplies. Noise problems are often exacerbated by improper grounding that fails to conduct noise away from the power system. In principle, noise consists of any unwanted distortion of the power signal that cannot be classified as harmonic distortion or transients. Noise disturbs electronic devices such as microcomputer and programmable controllers. The problem can be mitigated by using filters, isolation transformers, and line conditioners.

23.2.9 Voltage Fluctuation

Voltage fluctuations are systematic variations of the voltage envelope or a series of random voltage changes, the magnitude of which does not normally exceed the voltage ranges specified by ANSI C84.1 of 0.9 to 1.1 pu. IEC 61000-2-1 defines various types of voltage fluctuations. We will restrict our discussion here to IEC 61000-2-1 Type (d) voltage fluctuations, which are characterized as a series of random or continuous voltage fluctuations. Loads that can exhibit continuous, rapid variations in the load current magnitude can cause voltage variations that are often referred to as flicker. The term *flicker* is derived from the impact of the voltage fluctuation on lamps such that they are perceived to flicker by the human eye. To be technically correct, voltage fluctuation is an electromagnetic phenomenon while flicker is an undesirable result of the voltage fluctuation in some loads. However, the two terms are often linked together in standards. Therefore, we will also use the common term voltage flicker to describe such voltage fluctuations. Figure 23-9 illustrates a voltage waveform which produces flicker. This is caused by an arc furnace, one of the most common causes of voltage fluctuations on utility transmission and distribution systems.

23.2.10 Power Frequency Variations

Power frequency variations are defined as the deviation of the power system fundamental frequency from its specified nominal value (e.g., 50 or 60 Hz). The power system frequency is directly related to the rotational speed of the generators supplying the system. There are slight variations in frequency as the dynamic balance between load and generation changes. The size of the frequency shift and its duration depends on the load characteristics and the response of the generation control system

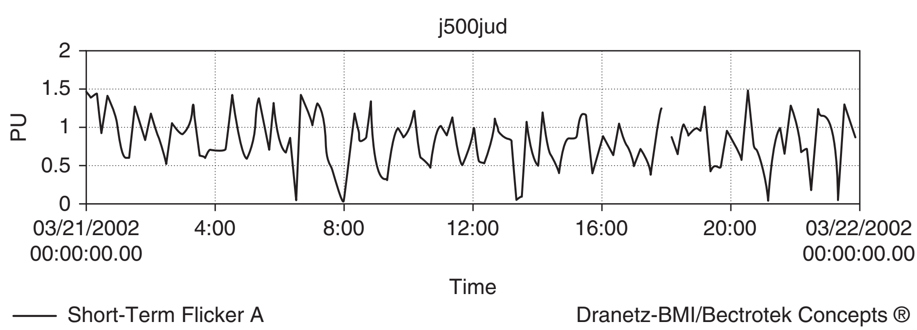


FIGURE 23-9 Flicker (Pst) at 161-kV substation bus measured according to IEC 61000-4-15.

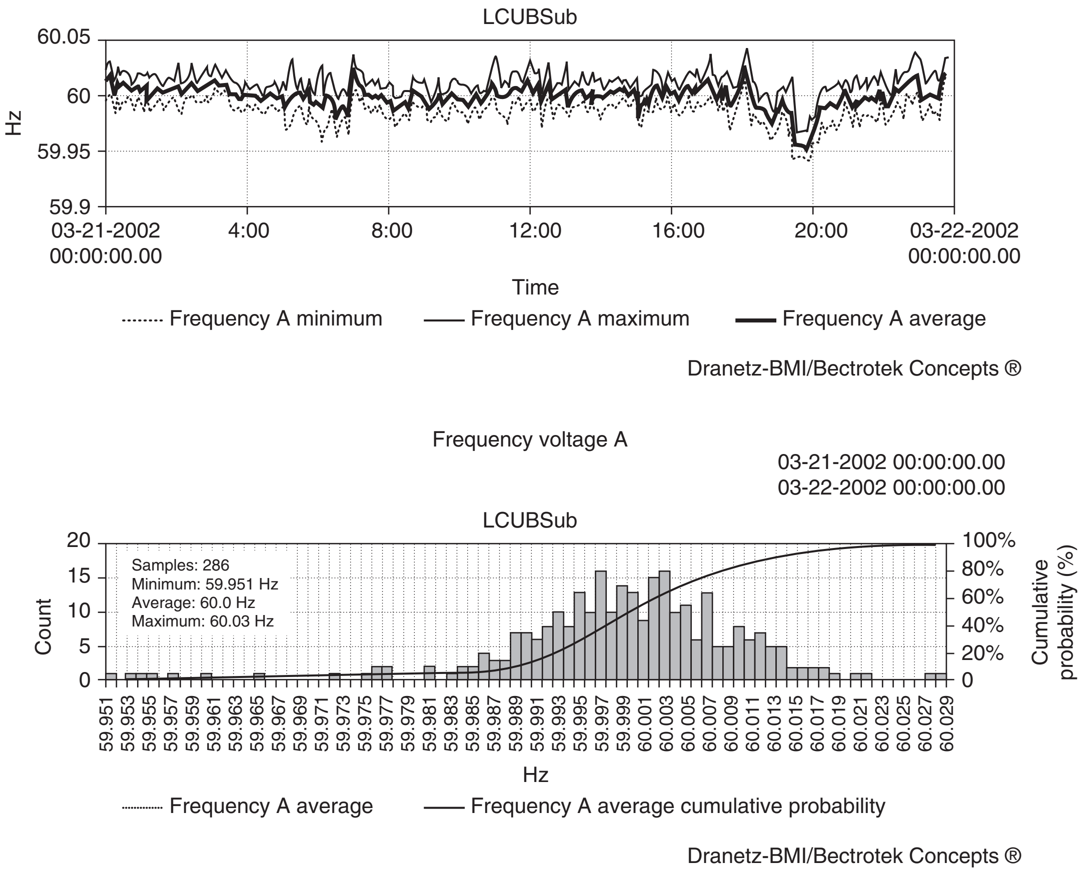


FIGURE 23-10 Power frequency trend and statistical distribution at 13-kV substation bus.

to load changes. Figure 23-10 illustrates frequency variations for a 24-h period on a typical 13-kV substation bus. Frequency variations that go outside of accepted limits for normal steady-state operation of the power system can be caused by faults on the bulk power transmission system, a large block of load being disconnected, or a large source of generation going offline.

23.3 VOLTAGE SAGS AND INTERRUPTIONS ON POWER SYSTEMS

23.3.1 Characteristics

Voltage sags and interruptions are related power quality problems. They are the result of faults in the power system and switching actions to isolate the faulted sections. A voltage sag is characterized by a short duration (typically 0.5 to 30 cycles) reduction in rms voltage caused by faults on the power system and the starting of large loads, such as motors. Momentary interruptions (typically no more than 2 to 5 s) cause a complete loss of voltage and are a common result of the actions taken by utilities to clear transient faults on their systems. Sustained interruptions of longer than 1 min are generally due to permanent faults. Due to the nature of the interconnected power systems and the utility fault-clearing schemes, voltage sags are the most common power quality disturbances.

23.3.2 Sources of Sags and Interruptions

Voltage sags and interruptions are generally caused by faults (short circuits) on the utility system and subsequent operations of protective devices in isolating the faults. Transient or temporary faults on the same or parallel feeders can result in voltage sags. Permanent faults usually result in interruptions. It is also possible that voltage sags are the result of starting of large loads, such as large motors. In some rare circumstances, energizing a transformer in a weak power system can also result in voltage sags. The voltage sag and interruption performance is greatly influenced by the utility feeder design and fault-clearing practices.

23.3.3 Utility System Fault Clearing

A radial distribution system is designed so that only one fault interrupter must operate to clear a fault. For permanent faults, that same device, or another, operates to *sectionalize* the feeder. That is, the faulted section is isolated so that power may be restored to the rest of the loads served from the sound sections. Orchestrating this process is referred to as the *coordination* of the overcurrent protection devices. While this is simple in concept, some of the behaviors of the devices involved can be quite complex. What is remarkable about this is that nearly all of the process is performed automatically by autonomous devices employing only local intelligence.

Overcurrent protection devices appear in series along a feeder. For permanent fault coordination, the devices operate progressively slower as one moves from the ends of the feeders toward the substation. This helps ensure the proper sectionalizing of the feeder so that only the faulted section is isolated. However, this principle is often violated for temporary faults, particularly if *fuse saving* is employed. The typical hierarchy of overcurrent protection devices on a feeder is

Feeder Breaker in the Substation. This is a circuit breaker capable of interrupting typically 40 kA of current and controlled by separate relays. When the available fault current is less than 20 kA, it is common to find reclosers used in this application.

Line Reclosers Mounted on Poles at Midfeeder. The simplest are self-contained with hydraulically-operated timing, interrupting, and reclosing mechanisms. Others have separate electronic controls.

Fuses on Many Lateral Taps Off the Main Feeder. These protective devices have significant implications on power quality issues.

23.3.4 Reclosers

Reclosers are a special circuit breaker designed to perform interruption and reclosing on temporary faults. They can reclose 2 or 3 times if needed in rapid succession. The multiple operations are designed to permit various sectionalizing schemes to operate and to give some more persistent transient faults a second chance to clear. The majority of faults will be cleared on the first operation. These devices can be found in numerous places along distribution feeders and sometimes in substations. They are typically applied at the head of sections subjected to numerous temporary faults. However, they may be applied nearly anywhere a convenient, low-cost primary-side circuit breaker is needed. Figure 23-11 shows a typical pole-mounted line recloser.

In addition to perform interruption and reclosing on temporary faults, reclosers are used for *fuse-saving or fast-tripping* applications. They are some of the fastest mechanical fault interrupters employed on the utility system. While they are typically rated for no faster than 3 to 6 cycles, many examples of interruptions as short as 1.5 cycles have been observed with power quality monitors.

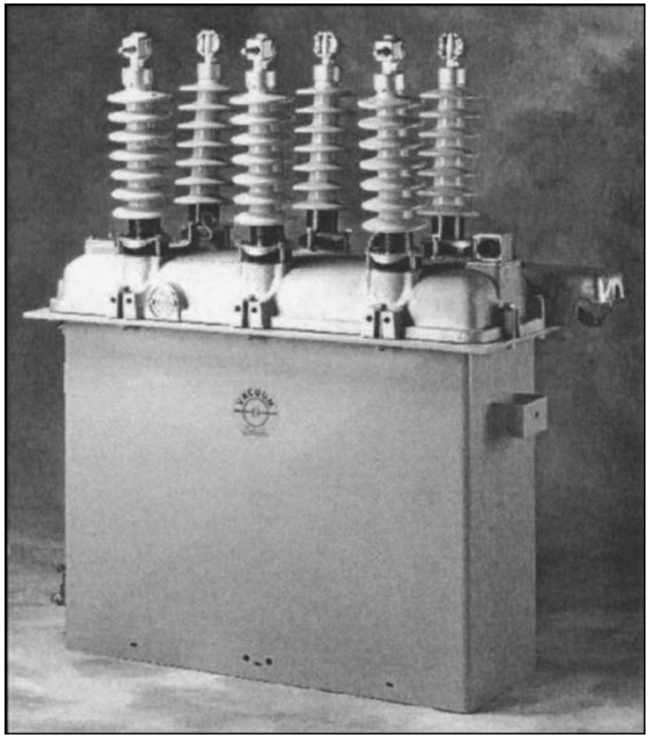


FIGURE 23-11 Typical standard 3-phase oil-insulated line recloser with vacuum interrupters. (*Photo courtesy of Cooper Power Systems.*)

This can be beneficial to limiting sag durations. Where fast tripping is not employed, the recloser control will commonly delay operation to more than 6 cycles to allow time for downline fuses to clear.

23.3.5 Reclosing Sequence

Reclosing is quite prevalent in North American utility systems. Utilities in regions of low lightning incidence may reclose only once because they assume that the majority of their faults will be permanent. In lightning-prone regions, it is common to attempt to clear the fault as many as 4 times. Figure 23-12 illustrates the two most common sequences in use on 4-shot reclosers:

1-fast operation, 3-delayed;

2-fast, 2-delayed.

Reclosers tend to have uniform reclose intervals between operations. The original hydraulic reclosers were limited to about 1 to 2 s and this setting has been retained by many utilities, although modern electronically controlled reclosers can be set for any value. It is common for the first reclose interval on some types of reclosers to be set for *instantaneous reclose*, which will result in closure in 12 to 30 cycles (0.2 to 0.5 s). This is done to reduce the time of the interruption and improve the power quality. However, there are some conflicts created by this, such as with distributed generation disconnecting times.

Substation circuit breakers often have a different style of reclosing sequence as shown in Fig. 23-13. This stems from a different evolution of relaying technology. Reclosing times are counted from the first tripping signal of the first operation. Thus, the common “0-15-45” operating sequence recloses

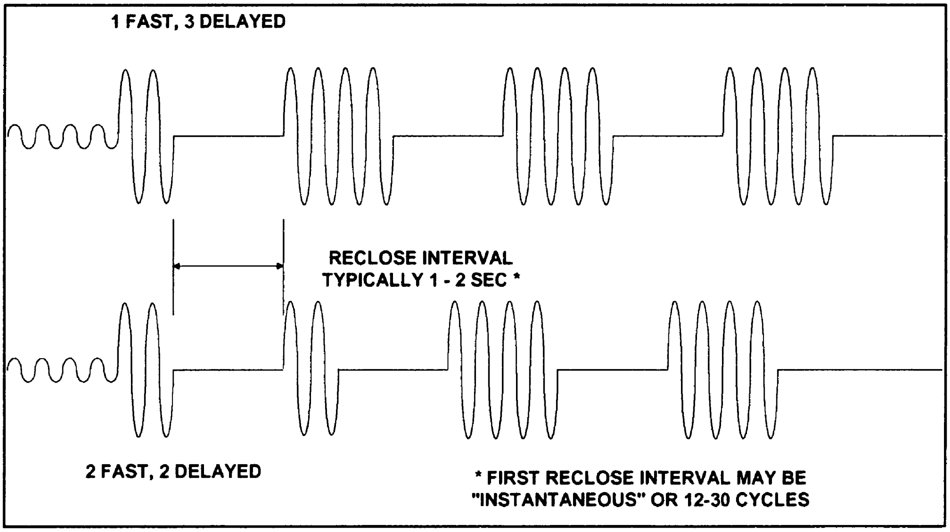


FIGURE 23-12 Common reclosing sequences for line reclosers in use in the United States.

essentially as fast as possible on the first operation, with approximately 15 and 30 s intervals between the next two operations.

Although the terminology may differ, modern breakers and reclosers can both be set to have the same operating sequences to meet load power quality requirements. Utilities generally choose one technology over the other based on cost or construction standards. It is generally fruitless to automatically reclose in distribution systems that are predominantly underground distribution cable, unless there is a significant portion that is overhead and exposed to trees or lightning.

23.3.6 Fuse Saving or Fast Tripping

Ideally, utility engineers would like to avoid blowing fuses needlessly on transient faults because a line crew must be dispatched to change it. Line reclosers were designed specifically to help save fuses. Substation circuit breakers can use instantaneous ground relaying to accomplish the same objective. The

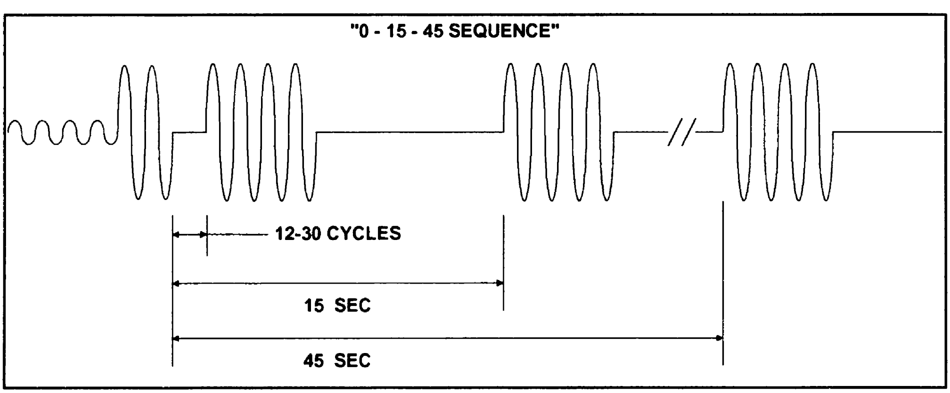


FIGURE 23-13 A common reclosing sequence for substation breakers in the United States.

basic idea is to have the mechanical circuit interrupting device operate very quickly on the first operation so that it clears before any fuses downline from it have a chance to melt. When the device closes back in, power is fully restored in the majority of the cases and no human intervention is required. The only inconvenience to the customer is a slight blink. This is called the fast operation of the device, or the instantaneous trip. If the fault is still there upon reclosing, there are two options in common usage:

1. Switch to a slow, or *delayed*, tripping characteristic. This is frequently the only option for substation circuit breakers; they will operate only one time on the instantaneous trip. This philosophy assumes that the fault is now permanent and switching to a delayed operation will give a downline fuse time to operate and clear the fault by isolating the faulted section.
2. Try a second fast operation. This philosophy is used where experience has shown a significant percentage of transient faults need two chances to clear while saving the fuses. Some line constructions and voltage levels have a greater likelihood that a lightning-induced arc may reignite and need a second chance to clear. Also, a certain percentage of tree faults will burn free if given a second shot.

Many utilities have abandoned fuse saving in selected areas due to complaints about power quality. The fast, or instantaneous, trip is eliminated so that breakers and reclosers have only time-delayed operations.

23.3.7 Fault-Induced Voltage Sags

The majority of voltage sags are caused by faults on the power systems and the subsequent operations of protective devices. Consider a customer that is supplied from the feeder supplied by circuit breaker no. 1 on the diagram shown in Fig. 23-14. If there is a fault on the same feeder, the customer will experience a voltage sag during the fault followed by an interruption when the breaker opens to clear the fault. If the fault is temporary in nature, a reclosing operation on the breaker should be successful and the interruption will only be temporary. It will usually require about 5 or 6 cycles for the breaker to operate, during which time a voltage sag occurs. The breaker will remain open for typically a minimum of 12 cycles up to 5 s depending on utility reclosing practices. Sensitive equipment will almost surely trip during this interruption.

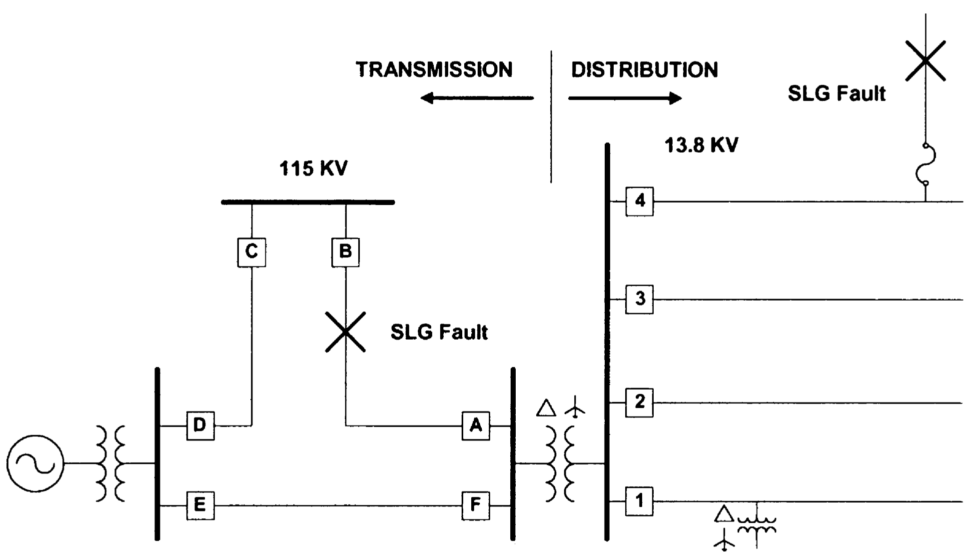


FIGURE 23-14 Fault locations on the utility power system.

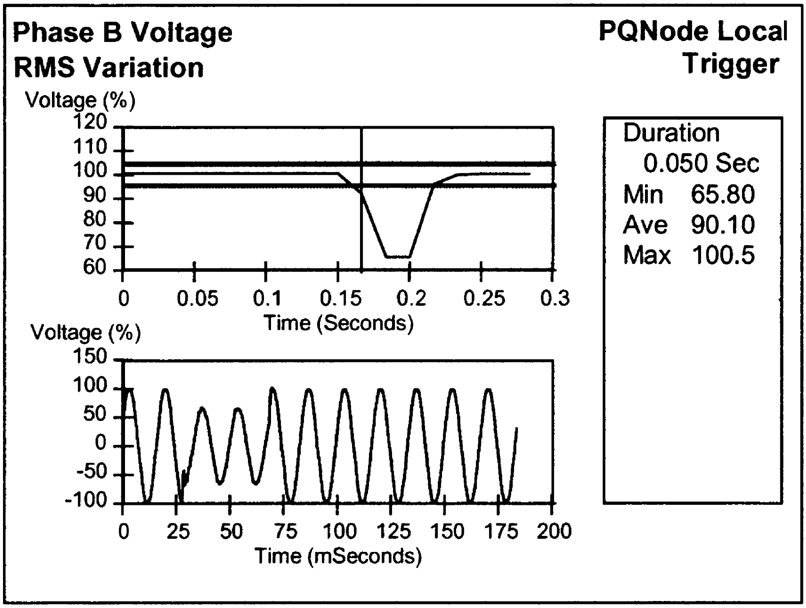


FIGURE 23-15 Voltage sag due to short circuit fault on a parallel utility feeder.

A much more common event would be a fault on one of the other feeders from the substation, that is, a fault is on a parallel feeder, or a fault somewhere on the transmission system (see the fault locations shown on Fig. 23-14). In either of these cases, the customer will experience a voltage sag during the period that the fault is actually on the system. As soon as breakers open to clear the fault, normal voltage will be restored at the customer. Note that to clear the fault shown on the transmission system, both breakers A and B must operate. Transmission breakers will typically clear a fault in 5 or 6 cycles. In this case there are two lines supplying the distribution substation and only one has a fault. Therefore, customers supplied from the substation should expect to see only a sag and not an interruption. The distribution fault on feeder no. 4 may be cleared either by the lateral fuse or the breaker, depending on the utility fuse saving practice. Any of these fault locations can cause equipment to misoperate in customer facilities. The relative importance of faults on the transmission system and the distribution system will depend on the specific characteristics of the systems (underground vs. overhead distribution, lightning flash densities, overhead exposure, etc.) and the sensitivity of the equipment to voltage sags.

Example of Voltage Sags due to a Fault on a Parallel Feeder. This example illustrates voltage sag and momentary interruption events due to a temporary fault on a utility feeder. Figures 23-15 and 23-16 show an interesting utility fault event recorded for an Electric Power Research Institute research project [13,14] by 8010 PQNode instruments* at two locations in the power system. The top chart in each of the figures is the rms voltage variation with time and the bottom chart is the first 175 ms of the actual waveform. Figure 23-15 shows the characteristic measured at a customer location on an unfaulted part of the feeder. Figure 23-16 shows the momentary interruption (actually two separate interruptions) observed downline from the fault. The interrupting device in this case was a line recloser that was able to interrupt the fault very quickly in about 2.5 cycles. This device can have a variety of settings. In this case, it was set for two fast operations and two delayed operations. Figure 23-15 shows only the brief sag to 65% voltage for the first fast operation. There was an

*PQNode is a registered trademark of Dranetz-BMI, Edison, NJ.

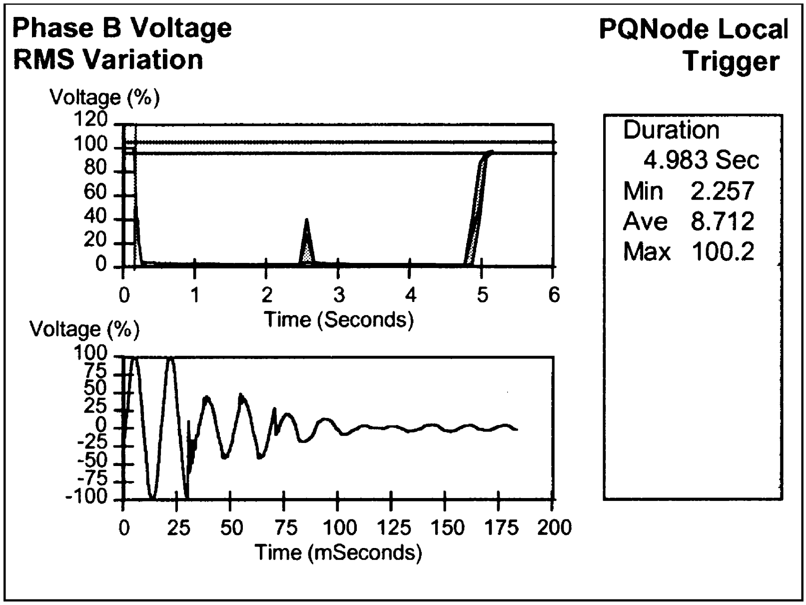


FIGURE 23-16 Utility short circuit fault event with two fast trip operations of utility line.

identical sag for the second operation. While this is very brief sag that is virtually unnoticeable by observing lighting blinks, many industrial processes would have shut down.

Figure 23-16 clearly shows the voltage sag prior to fault clearing and the subsequent two fast recloser operations. The reclose time (the time the recloser was open) was a little more than 2 s, a very common time for a utility line recloser. Apparently, the fault—perhaps, a tree branch—was not cleared completely by the first operation, forcing a second. The system was restored after the second operation.

23.3.8 Motor Starting Sags

Motors have the undesirable effect of drawing several times their full load current while starting. This large current will, by flowing through system impedances, cause a voltage sag which may dim lights, cause contactors to drop out, and disrupt sensitive equipment. The situation is made worse by an extremely poor starting displacement factor—usually in the range of 15%, 30%. The time required for the motor to accelerate to rated speed increases with the magnitude of the sag, and an excessive sag may prevent the motor from starting successfully. Motor starting sags can persist for many seconds, as illustrated in Fig. 23-17.

23.3.9 Motor Starting Methods

Energizing the motor in a single step (*full voltage starting*) provides low cost and allows the most rapid acceleration. It is the preferred method unless the resulting voltage sag or mechanical stress is excessive.

Autotransformer starters have two autotransformers connected in open delta. Taps provide a motor voltage of 80%, 65%, or 50% of system voltage during start-up. Line current and starting torque vary with the square of the voltage applied to the motor, so the 50% tap will deliver only 25% of the full voltage starting current and torque. The lowest tap which will supply the required starting torque is selected.

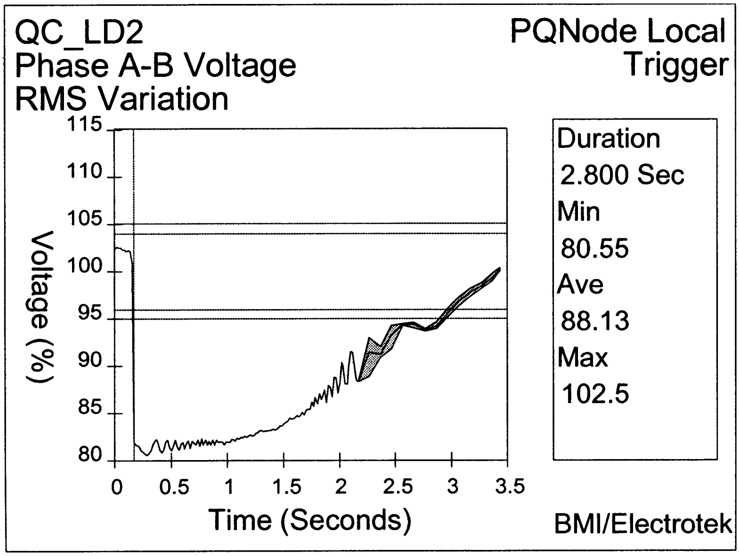


FIGURE 23-17 Typical motor starting voltage sag.

Resistance and reactance starters initially insert an impedance in series with the motor. After a time delay, this impedance is shorted out. Starting resistors may be shorted out over several steps; starting reactors are shorted out in a single step. Line current and starting torque vary directly with the voltage applied to the motor, so for a given starting voltage, these starters draw more current from the line than with autotransformer starters, but provide higher starting torque. Reactors are typically provided with 50%, 45%, and 37.5% taps.

Part winding starters are attractive for use with dual-rated motors (220/440 or 230/460V). The stator of a dual-rated motor consists of two windings connected in parallel at the lower voltage rating, or in series at the higher voltage rating. When operated with a part winding starter at the lower voltage rating, only one winding is energized initially, limiting starting current and starting torque to 50% of the values seen when both windings are energized simultaneously.

Wye-Delta starters connect the stator in wye for starting, then after a time delay, reconnect the windings in delta. The wye connecting reduces the starting voltage to 57% of the system line-line voltage; starting current and starting torque are reduced to 33% of their values for full voltage start.

23.3.10 Estimating the Sag Severity during Full Voltage Starting

As shown in Fig. 23-17, starting an induction motor results in a steep dip in voltage, followed by a gradual recovery. If full voltage starting is used, the sag voltage, in per unit of nominal system voltage is

$$V_{min}(pu) = \frac{V(pu) \cdot kVA_{SC}}{kVA_{LR} + kVA_{SC}} \tag{23-1}$$

where $V(pu)$ = is the actual system voltage, in per unit of nominal
kVA_{LR} = is the motor locked rotor kVA
kVA_{SC} = is the system short circuit kVA at the motor

Figure 23-18 illustrates the results of this computation for sag to 90% of nominal voltage, using typical system impedances and motor characteristics.

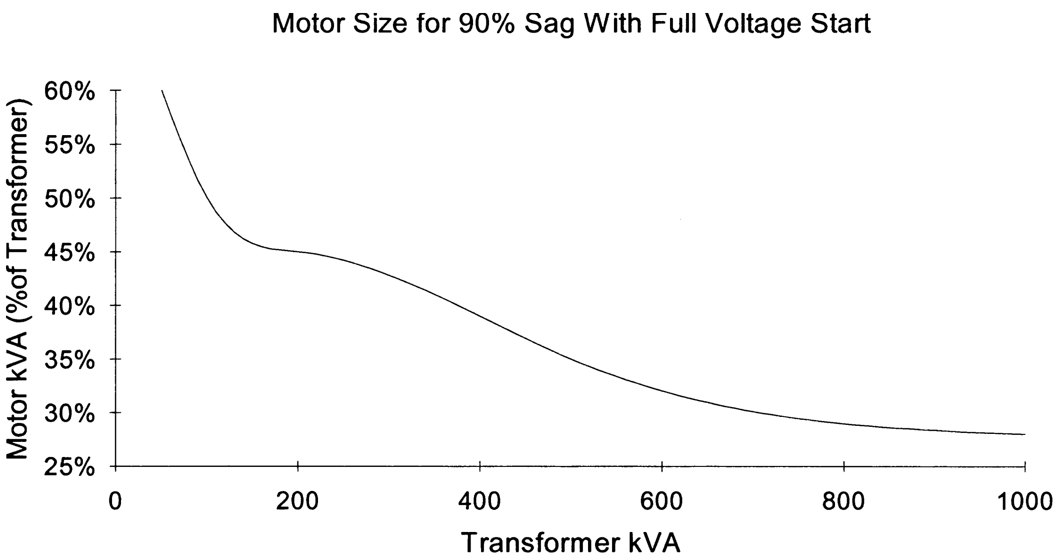


FIGURE 23-18 Typical motor vs. transformer size for full voltage starting sags of 90%.

If the result is above the minimum allowable steady-state voltage for the effected equipment, then the full voltage starting is acceptable. If not, then the sag magnitude versus duration characteristic must be compared to the voltage tolerance envelope of the effected equipment. The required calculations are fairly complicated, and best left to a motor starting or general transient analysis computer program.

23.4 ELECTRICAL TRANSIENT PHENOMENA

23.4.1 Sources and Characteristics

In principle, electrical transient phenomena can be generated due to natural events such as lightning strokes, and switching operations such as capacitor, load, and transformer energizing, and protective device operations. However, two main sources of transient overvoltages on utility systems are capacitor switching and lightning.

23.4.2 Capacitor Switching Transient Overvoltages

Capacitor switching is one of the most common switching events on utility systems. Capacitors are used to provide reactive power (vars) to correct the power factor, which reduces losses and supports the voltage on the system. One drawback to capacitors is that they yield oscillatory transients when switched. Some capacitors are energized all the time (a fixed bank) while others are switched according to load levels. Various control means are used to determine when they are switched including time, temperature, voltage, current, and reactive power. It is common for controls to combine two or more of these functions, such as temperature with voltage override.

Figure 23-19 shows the one-line diagram of a typical utility feeder capacitor switching situation. When the switch is closed, a transient similar to the one in Fig. 23-20 may be observed upline from the capacitor at the monitor location. In this particular case, the capacitor switch contacts close at a point near the system voltage peak. This is common for many types of switches because the insulation across the switch contacts tends to break down when the voltage across the switch is at a maximum

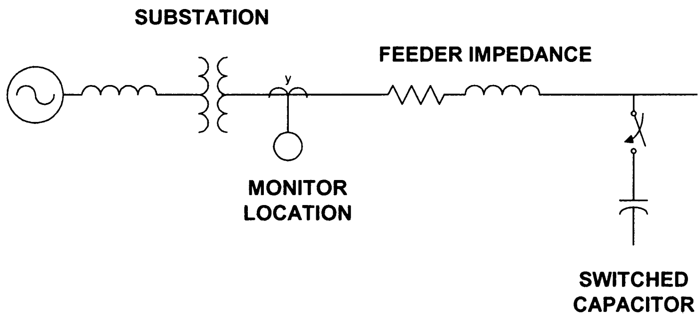


FIGURE 23-19 One-line diagram of capacitor switching operation.

value. The voltage across the capacitor at this instant is zero. Since the capacitor voltage cannot change instantaneously, the system voltage at the capacitor location is briefly pulled down to zero and rises as the capacitor begins to charge toward the system voltage. Because the power system source is inductive, the capacitor voltage overshoots and rings at the natural frequency of the system. At the monitoring location shown, the initial change in voltage will not go completely to zero because of the impedance between the observation point and the switched capacitor. However, the initial drop and subsequent ringing transient that is indicative of a capacitor switching event will be observable to some degree. The overshoot will generate a transient between 1.0 and 2.0 pu depending

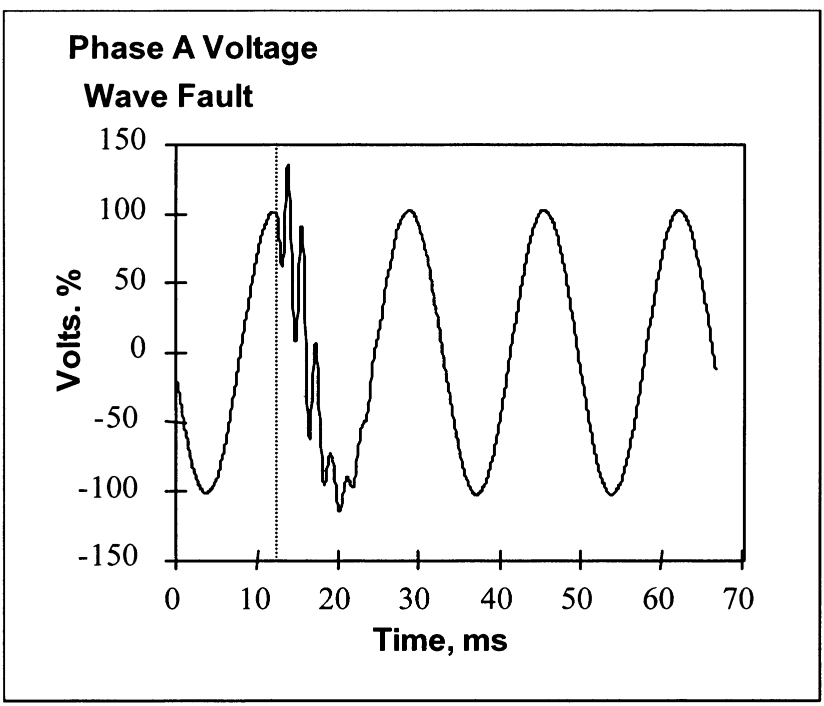


FIGURE 23-20 Typical utility capacitor switching transient reaching 134% voltage, observed upline from the capacitor.

on system damping. In this case, the transient observed at the monitoring location is about 1.34 pu. Utility capacitor switching transients are commonly in the 1.3 to 1.4 pu range, but have also been observed near the theoretical maximum.

23.4.3 Magnification of Capacitor Switching Transient Overvoltages

Capacitor switching transients can propagate into the local power system and will generally pass through distribution transformers into customer load facilities by nearly the amount related to the turns ratio of the transformer. If there are capacitors on the secondary system, the voltage may actually be magnified on the load side of the transformer if the natural frequencies of the systems are properly aligned. The circuit of concern for this phenomenon is illustrated in Fig. 23-21. Transient overvoltages on the end-user side may reach as high as 3.0 to 4.0 pu on the low-voltage bus under these conditions, with potentially damaging consequences for all types of customer equipment.

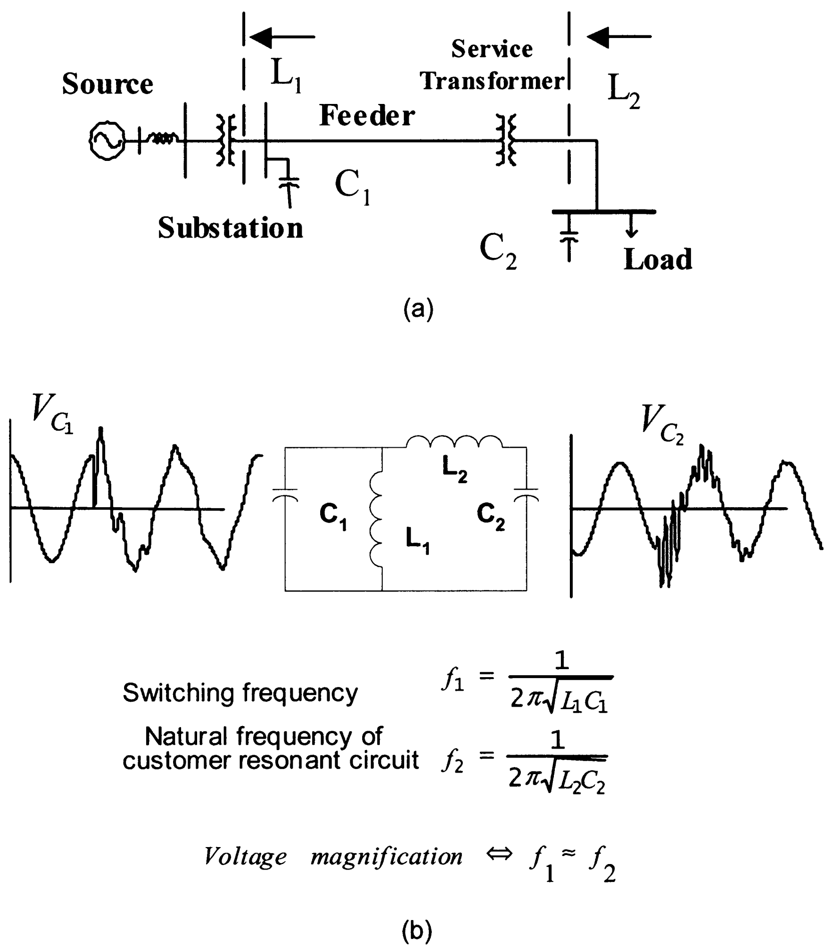


FIGURE 23-21 Voltage magnification of capacitor bank switching. (*a*) Voltage magnification at customer capacitor due to energizing capacitor on utility system (*b*) Equivalent circuit

23.4.4 Options to Limit Magnification

Magnification of utility capacitor switching transients at the end-user location occurs over a wide range of transformer and capacitor sizes. Resizing the customer's power factor correction capacitors or step-down transformer is therefore usually not a practical solution. One solution is to control the transient overvoltage at the utility capacitor. This is sometimes possible using synchronous closing breakers or switches with preinsertion resistors. At the customer location, high-energy surge arresters can be applied to limit the transient voltage magnitude at the customer bus. Energy levels associated with the magnified transient will typically be in the range of 1 kJ. Figure 23-22 shows the expected arrester energy for a range of low-voltage capacitor sizes. High energy MOV arresters for low-voltage applications can withstand 2 to 4 kJ.

While such brief transients up to 2.0 per unit are not generally damaging to the system insulation, it can often cause misoperation of electronic power conversion devices. Controllers may interpret the high voltage as a sign that there is an impending dangerous situation and subsequently disconnect the load to be safe. The transient may also interfere with the gating of thyristors. It is important to note that the arresters can only limit the transient to the arrester protective level. This will typically be approximately 1.8 times the normal peak voltage (1.8 pu).

Another means of limiting the voltage magnification transient is to convert the end-user, power-factor-correction banks to harmonic filters. An inductance in series with the power-factor-correction bank will decrease the transient voltage at the customer bus to acceptable levels. This solution has multiple benefits by providing correction for displacement power factor, controlling harmonic distortion levels within the facility, and limiting the concern for magnified capacitor switching transients.

In many cases, there are only a small number of load devices, such as adjustable-speed motor drives, that are adversely affected by the transient. It is frequently more economical to place line reactors in series with the drives to block the high frequency magnification transient. A 3% reactor is generally effective. While offering only a small impedance to power frequency current, it offers a considerably larger impedance to the transient. Many types of drives have this protection inherently, either through an isolation transformer or a dc bus reactance.

23.4.5 Options to Limit Capacitor Switching Transients—Preinsertion

Preinsertion resistors can reduce the capacitor switching transient considerably. The first peak of the transient is usually the most damaging. The idea is to insert a resistor into the circuit briefly so that the first

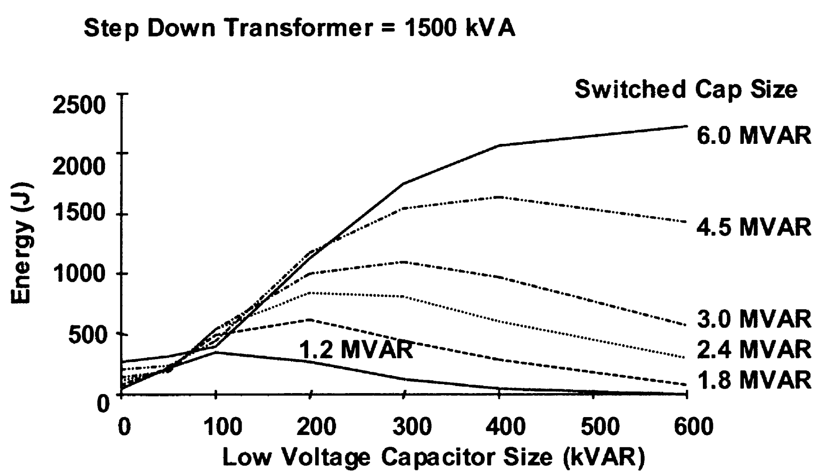


FIGURE 23-22 Arrester energy duty caused by magnified transient.

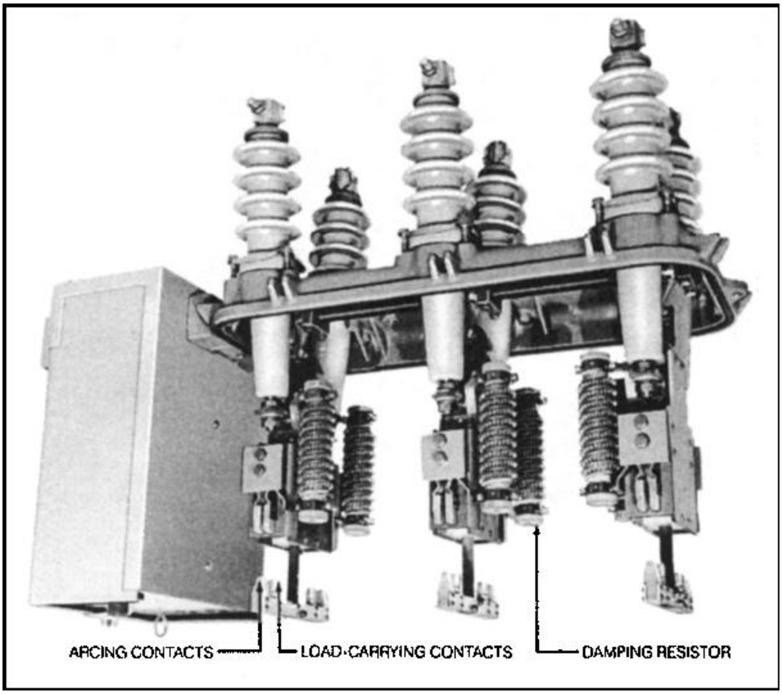


FIGURE 23-23 Capacitor switch with preinsertion resistors. (*Courtesy of Cooper Power Systems.*)

peak is damped significantly. This is old technology, but still quite effective. Figure 23-23 shows one example of a capacitor switch with preinsertion resistors to reduce transients. The preinsertion is accomplished by the movable contacts sliding past the resistor contacts first before mating with the main contacts. This results in a preinsertion time of approximately one-fourth of a cycle at 60 Hz. The effectiveness of the resistors is dependent on capacitor size and available short-circuit current at the capacitor location. Table 23-3 shows expected maximum transient overvoltages upon energization for various conditions, both with and without the preinsertion resistors. These are the maximum values expected; average values are typically 1.3 to 1.4 pu without resistors and 1.1. to 1.2 with resistors.

TABLE 23-3 Peak Transient Overvoltages Due to Capacitor Switching With and Without Preinsertion Resistor

Size, kvar	Avail. Short Circuit, kA	Without Resistor (pu)	With 6.4 Ω Resistor (pu)
900	4	1.95	1.55
900	9	1.97	1.45
900	14	1.98	1.39
1200	4	1.94	1.50
1200	9	1.97	1.40
1200	14	1.98	1.34
1800	4	1.92	1.42
1800	9	1.96	1.33
1800	14	1.97	1.28

Courtesy of Cooper Power Systems

23.4.6 Options to Limit Capacitor Transient Switching—Synchronous Closing

Another popular strategy for reducing transients on capacitor switching is to use a synchronous closing breaker. This is a relatively new technology for controlling capacitor switching transients. Synchronous closing prevents transients by timing the contact closure such that the system voltage closely matches the capacitor voltage at the instant the contacts make. This avoids the step change in voltage that normally occurs when capacitors are switched, causing the circuit to oscillate. Figure 23-24 shows a vacuum switch made for this purpose. It is applied on 46-kV-class capacitor banks. It consists of three independent poles with separate controls. The timing for synchronous closing is determined by anticipating an upcoming voltage zero. Its success is dependent on the consistent operation of the vacuum switch. The switch reduces capacitor inrush currents by an order of magnitude and voltage transients to about 1.1 pu. A similar switch may also be used at distribution voltages. Each of the switches described here requires a sophisticated microprocessor-based control. Understandably, a synchronous closing system is more expensive than a straightforward capacitor switch. However, it is frequently a cost-effective solution when capacitor switching transients are disrupting end-user loads.

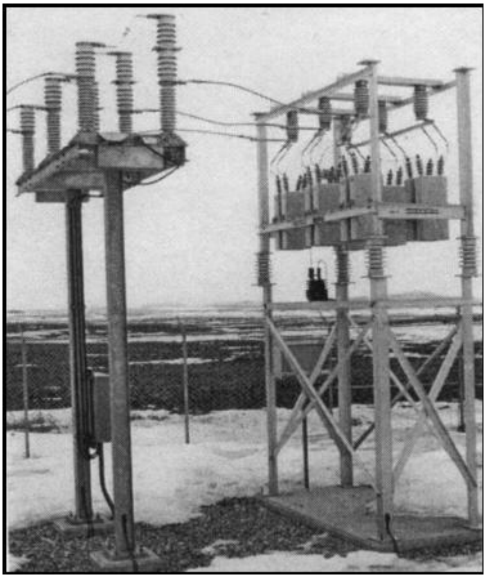


FIGURE 23-24 Synchronous closing capacitor switch. (*Courtesy of Joslyn Hi-Voltage Corporation.*)

23.4.7 Lightning

Lightning is a potent source of impulsive transients and can have serious impacts on power system and end-user equipment. Figure 23-25 illustrates some of the places where lightning can strike that results in lightning currents being conducted from the power system into loads. The most obvious conduction path occurs during a direct strike to a phase wire, either on the primary or the secondary side of the

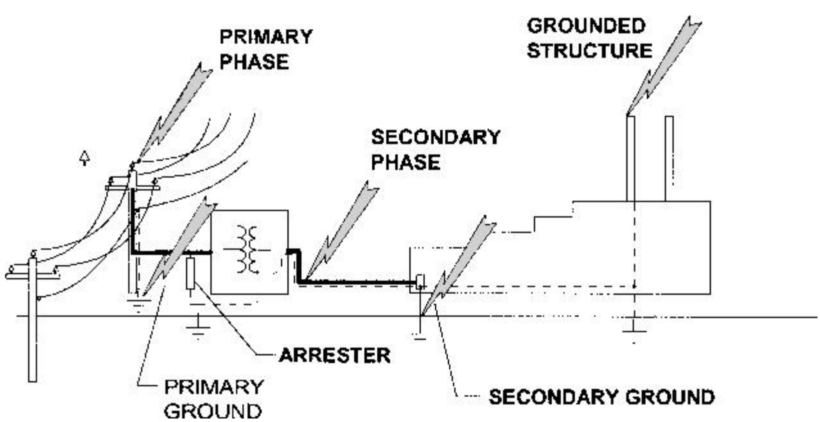


FIGURE 23-25 Stroke locations for conduction of lightning impulses into load facilities.

transformer. This can generate very high overvoltages, but some analysts question whether this is the most common way that lightning surges enter load facilities and cause damage. Very similar transient overvoltages can be generated by lightning currents flowing along ground conductor paths. Note that there can be numerous paths for lightning currents to enter the grounding system. Common ones, indicated by the dotted lines in Fig. 23-25, include the primary ground, the secondary ground, and the structure of the load facilities. Note also that strokes to the primary phase are conducted to the ground circuits through the arresters on the service transformer. Thus, many more lightning impulses may be observed at loads than one might think. Note that grounds are never perfect conductors, especially for impulses. While most of the surge current may eventually be dissipated into the ground connection closest to the stroke, there will be substantial surge currents flowing in other connected ground conductors in the first few microseconds of the strike.

The chief power quality problems with lightning stroke currents entering the ground system are

- They raise the potential of the local ground above other grounds in the vicinity by several kilovolts. Sensitive electronic equipment that is connected between two ground references, such as a computer connected to the telephone system through a modem, can fail when subjected to the lightning surge voltages.
- They induce high voltages in phase conductors as they pass through cables on the way to a better ground.

23.4.8 Low-Side Surges

Some utility and end-user problems with lightning impulses are closely related. One of the most significant ones is called the *low-side surge* problem by many utility engineers. The name was coined by distribution transformer designers because it appears from the transformer's perspective that a current surge is suddenly injected into the low-voltage side terminals. Utilities have not applied secondary arresters at low-voltage levels in great numbers. From the customer's point of view, it appears to be an impulse coming from the utility and is likely to be termed as "secondary surge."

Both problems actually have different side effects of the same surge phenomenon—lightning current flowing from either the utility side or the customer side along the service cable neutral. Figure 23-26 shows one possible scenario. Lightning strikes the primary line and the current is discharged through

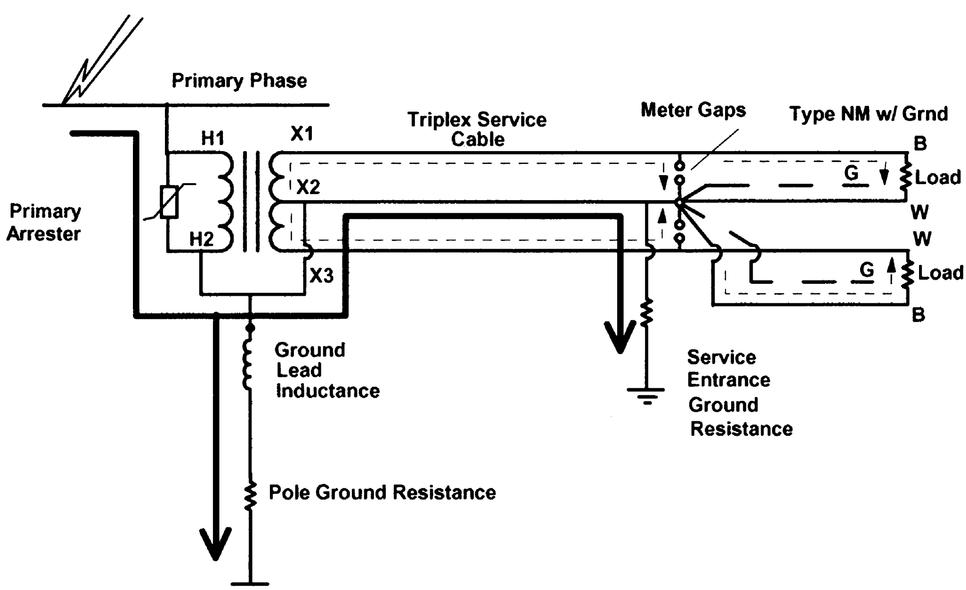


FIGURE 23-26 Primary arrester discharge current divides between pole and load ground.

the primary arrester to the pole ground lead. This lead is also connected to the X2 bushing of the transformer at the top of the pole. Thus, some of the current will flow toward the load ground. The amount of current into the load ground is primarily dependent on the size of the pole ground resistance relative to the load ground. Inductive elements may play a significant role in the current division for the front of the surge, but the ground resistances basically dictate the division of the bulk of the stroke current.

The current that flows through the secondary cables causes a voltage drop in the neutral conductor that is only partially compensated by mutual inductive effects with the phase conductors. Thus, there is a net voltage across the cable, forcing current through the transformer secondary windings and into the load as shown by the dashed lines in the figure. If there is a complete path, substantial surge current will flow. As it flows through the transformer secondary, a surge voltage is induced in the primary, sometimes causing a layer-to-layer insulation failure near the grounded end. If there is not a complete path, the voltage will buildup across the load and may flash over somewhere on the secondary. It is common for the meter gaps to flashover, but not always before there is damage on the secondary because the meter gaps are usually 6 to 8 kV, or higher. The amount of voltage induced in the cable is dependent on the rate-of-rise of the current, which is dependent on other circuit parameters as well as the lightning stroke.

The chief power quality problems this causes are

- The impulse entering the load can cause failure or misoperation of load equipment.
- The utility transformer will fail causing an extended power outage.
- The failing transformer may subject the load to sustained steady-state overvoltages because part of the primary winding is shorted, decreasing the transformer turns ratio. Failure usually occurs in seconds, but has been known to take hours.

The key to this problem is the amount of surge current traveling through the secondary service cable. Keep in mind that the same effect occurs regardless of the direction of the current. All that is required is for the current to get into the ground circuits and for a substantial portion to flow through the cable on its way to another ground. Thus, lightning strikes to either the utility system or the end-user facilities can produce the same symptoms. Transformer protection is more of an issue in residential services, but the secondary transients will appear in industrial systems as well.

23.4.9 Low-Side Surges—An Example

Figure 23-27 shows a waveform of the open-circuit voltage measured at an electrical outlet location in a laboratory mock-up of a residential service [16]. For a relatively small stroke to the primary line (2.6 kA), the voltages at the outlet reached nearly 15 kV. In fact, higher current strokes caused random flashovers of the test circuit, which made measurements difficult. This reported experience is indicative of the capacity of these surges to cause overvoltage problems.

The waveform is a very high-frequency, ringing wave riding on the main part of the low-side surge. The ringing is very sensitive to the cable lengths. A small amount of resistive load, such as a lightbulb, would contribute greatly to the damping. The ringing wave differs depending on where the surge was applied while the base low-side surge wave remains about the same; it is more dependent on the waveform of the current through the service cable. One interesting aspect of this wave is that the ringing is so fast that it gets by the spark gaps in the meter base even though the voltage is 2 times the nominal sparkover value. In the tests, the outlets and lamp sockets could also withstand this kind of wave for about 1 μs before they flashed over. Thus, it is possible to have some high overvoltages propagating throughout the system. The waveform in this figure represents the available open-circuit voltage. In actual practice, a flashover would have occurred somewhere in the circuit after a brief time.

23.4.10 Ferroresonance

The term *ferroresonance* refers to a resonance that involves capacitance and iron-core inductance. The most common condition in which it causes disturbances in the power system is when the magnetizing

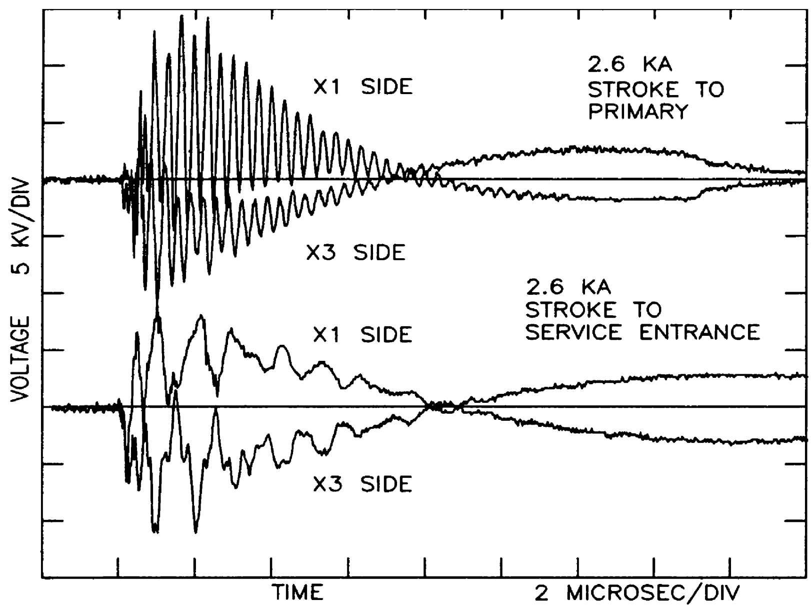
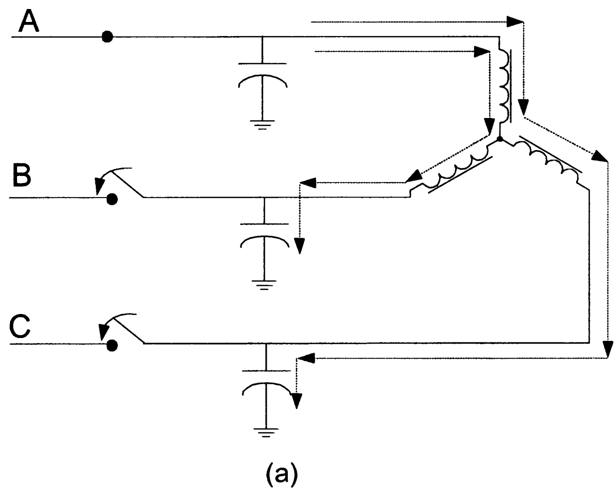


FIGURE 23-27 Voltage appearing at outlet due to low-side surge phenomena.

impedance of a transformer is placed in series with a system capacitor due to an open-phase conductor. Under controlled conditions, ferroresonance can be exploited for useful purpose such as in a constant-voltage transformer. In practice, ferroresonance most commonly occurs when unloaded transformers become isolated on underground cables of a certain range of lengths. The capacitance of overhead distribution lines is generally insufficient to yield the appropriate conditions.

The minimum length of cable required to cause ferroresonance varies with system voltage level. The capacitance of cables is nearly the same for all distribution voltage levels, varying from 40 to 100 nF per 1000 ft, depending on conductor size. However, the magnetizing reactance of a 35-kV-class distribution transformer is several times higher (curve is steeper) than a comparably-sized 15-kV-class transformer. Therefore, damaging ferroresonance has been more common at the higher voltages. For delta-connected transformers, ferroresonance can occur for less than 100 ft of cable. For this reason, many utilities avoid this connection on cable-fed transformers. The grounded wye-wye transformer has become the most commonly used connection in underground systems in North America. It is more resistant, but not immune, to ferroresonance because most units use a three-legged or five-legged core design that couples the phases magnetically. It may require a minimum of several hundred feet of cable to provide enough capacitance to create a ferroresonant condition for this connection. The most common events leading to ferroresonance are

- Manual switching of an unloaded, cable-fed, 3-phase transformer where only one phase is closed (Fig. 23-28*a*). Ferroresonance may be noted when the first phase is closed upon energization or before the last phase is opened on de-energization.
- Manual switching of an unloaded, cable-fed, 3-phase transformer where one of the phases is open (Fig. 23-28*b*). Again, this may happen during energization or de-energization.
- One or two riser-pole fuses may blow leaving a transformer with one or two phases open. Single-phase reclosers may also cause this condition. Today, many modern commercial loads will have controls that transfer the load to backup systems when they sense this condition. Unfortunately, this leaves the transformer without any load to damp out the resonance.
- Phase of a cable connected to a wye-connected transformer.



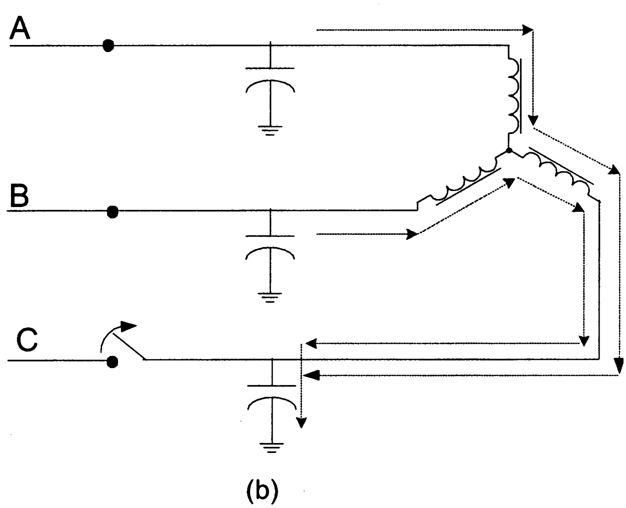


FIGURE 23-28 Common system conditions where ferroresonance may occur: (*a*) one phase closed, (*b*) one phase open.

23.4.11 Transformer Energizing

Energizing a transformer produces inrush currents that are rich in harmonic components for a period lasting up to 1 s. If the system has a parallel resonance near one of the harmonic frequencies, a dynamic overvoltage condition results that can cause failure of arresters and problems with sensitive equipment. This problem can occur when large transformers are energized simultaneously with large power factor correction capacitor banks in industrial facilities. The equivalent circuit is shown in Fig. 23-29. A dynamic overvoltage waveform caused by a third-harmonic resonance in the circuit is shown in Fig. 23-30. After the expected

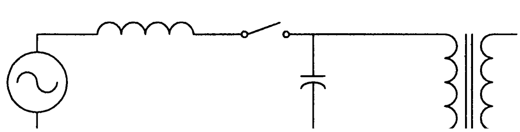


FIGURE 23-29 Energizing a capacitor and transformer simultaneously can lead to dynamic overvoltages.

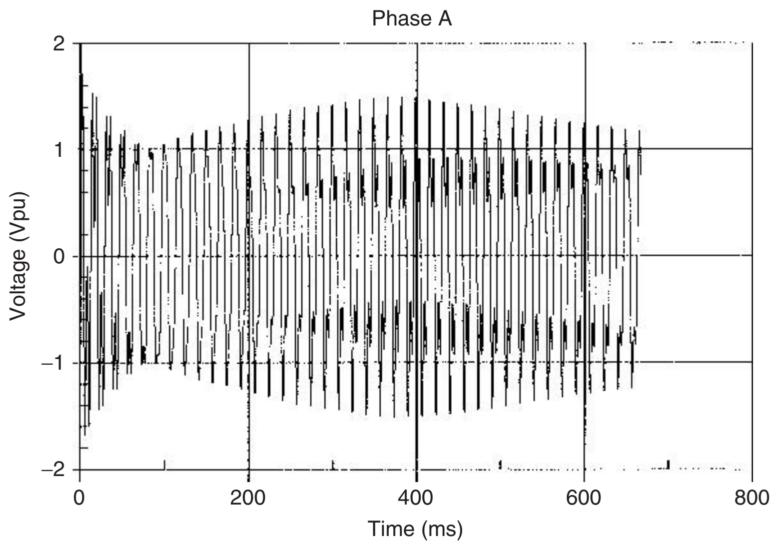


FIGURE 23-30 Dynamic overvoltages during transformer energizing.

initial transient, the voltage again swells to nearly 150% for many cycles until the losses and load damp out the oscillations. This can place severe stress on some arresters and has been known to significantly shorten the life of capacitors.

This form of dynamic overvoltage problem can often be eliminated simply by not energizing the capacitor and transformer together. One plant solved the problem by energizing the transformer first and not energizing the capacitor until load was about to be connected to the transformer.

23.5 POWER SYSTEMS HARMONICS

23.5.1 General

Harmonic distortion is not a new phenomenon on power systems. Concern over distortion has ebbed and flowed a number of times during the history of ac electric power systems. Scanning the technical literature of the 1930s and 1940s, one will notice many articles on the subject. Then the primary sources were the transformers and the primary problem was inductive interference with open-wire telephone systems. The forerunners of modern arc lighting were being introduced and were causing quite a stir because of their harmonic content—not unlike the stir caused by electronic power converters in more recent times.

In contrast, voltage sags and interruptions are nearly universal to every feeder and represent the most numerous and significant power quality deviations. The end user sector suffers more from harmonic problems than the utility sector. Industrial users with adjustable speed drives, arc furnaces, induction furnaces, and the like, are much more susceptible to problems stemming from harmonic distortion.

A good assumption for most utilities in the United States is that the sine wave voltage generated in central power stations is very good. In most areas, the voltage found on transmission systems typically has much less than 1% distortion. However, the distortion increases closer to the load. At some loads, the current waveforms barely resemble a sine wave. Electronic power converters can chop the current into seemingly arbitrary waveforms.

23.5.2 Harmonic Distortion

Harmonic distortion is caused by nonlinear devices in the power system. A nonlinear device is one in which the current is not proportional to the applied voltage. Figure 23-31 illustrates this concept by the case of a sinusoidal voltage applied to a simple nonlinear resistor in which the voltage and current vary according to the curve shown. While the applied voltage is perfectly sinusoidal, the resulting current is distorted. Increasing the voltage by a few percent may cause the current to double and take on a different waveshape. This is the source of most harmonic distortion in a power system. Figure 23-32 illustrates that any periodic, distorted waveform can be expressed as a sum of sinusoids. When a waveform is identical from one cycle to the next, it can be represented as a sum of pure sine waves in which the frequency of each sinusoid is an integer multiple of the fundamental frequency of the distorted wave. This multiple is called a *harmonic* of the fundamental, hence the name of this subject matter. The sum of sinusoids is referred to as a *Fourier series*, named after the great mathematician who discovered the concept.

23.5.3 Voltage and Current Distortion

The term "harmonics" is often used by itself without further qualification. Generally, it could mean one of the following three:

1. The harmonic voltages are too great (the voltage too distorted) for the control to properly determine firing angles.
2. The harmonic currents are too great for the capacity of some device in the power supply system such as a transformer and the machine must be operated at a lower than rated power.
3. The harmonic voltages are too great because the harmonic currents produced by the device are too great for the given system condition.

Clearly, there are separate causes and effects for voltages and currents as well as some relationship between them. Thus, the term harmonics by itself is inadequate to definitively describe a problem. Nonlinear loads appear to be sources of harmonic current in shunt with and injecting

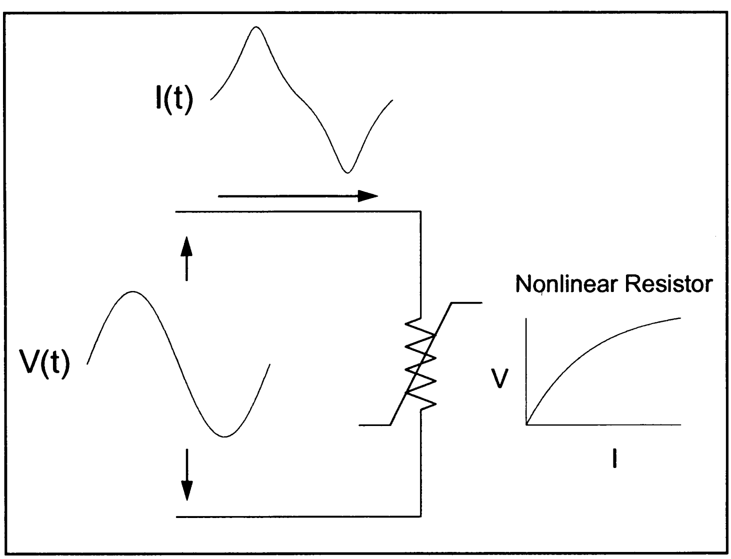


FIGURE 23-31 Current distortion caused by nonlinear resistance.

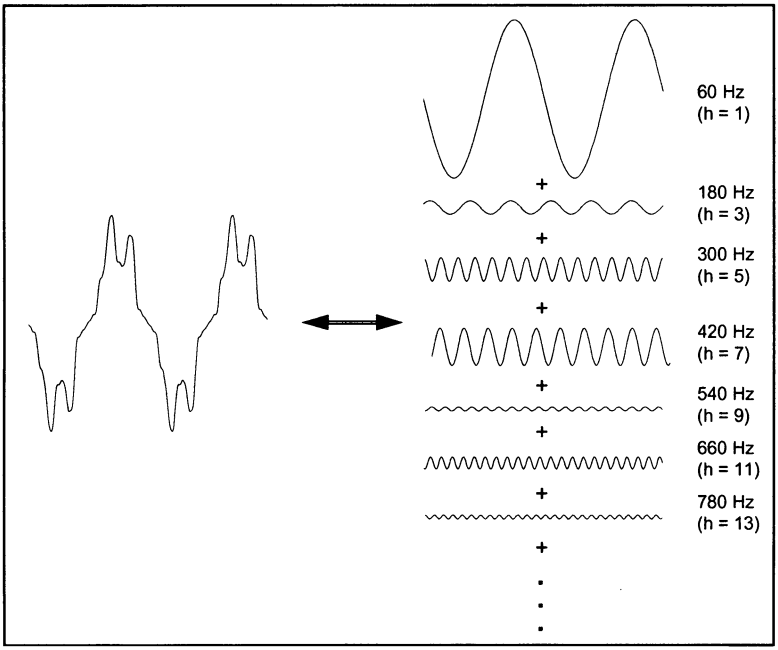


FIGURE 23-32 Fourier series representation of a distorted waveform.

harmonic currents into the power system. For nearly all analyses, it is sufficient to treat these harmonic-producing loads simply as current sources. There are exceptions to this as described later.

Voltage distortion is the result of distorted currents passing through the linear, series impedance of the power delivery system as illustrated in Fig. 23-33. Although, assuming that the source bus is ultimately a pure sinusoid, there is a nonlinear load that draws a distorted current. The harmonic currents passing through the impedance of the system cause a voltage drop for each harmonic. This results in voltage harmonics appearing at the load bus. The amount of voltage distortion depends on the impedance and the current. Assuming the load bus distortion stays within reasonable limits (e.g., less than 5%), the amount of harmonic current produced by the load is generally constant.

While the load current harmonics ultimately cause the voltage distortion, it should be noted that load has no control over the voltage distortion. The same load put in two different locations on the power system will result in two different voltage distortion values. Recognition of this fact is the

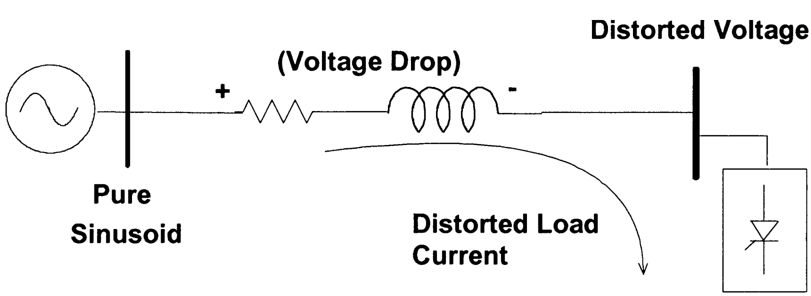


FIGURE 23-33 Harmonic currents flowing through the system impedance results in harmonic voltages at the load.

basis for the division of responsibilities for harmonic control that are found in standards such as IEEE Std 519-1992:

- The control over the amount of harmonic current injected into the system takes place at the end-use application,
- Assuming the harmonic current injection is within reasonable limits, the control over the voltage distortion is exercised by the entity having control over the system impedance, which is often the utility.

One must be careful when describing harmonic phenomena to understand that there are distinct differences between the causes and effects of harmonic voltages and currents. The use of the term harmonics should be qualified accordingly. By popular convention in the power industry, the majority of time the term is used by itself when referring to load apparatus, the speaker is referring to the harmonic currents. When referring to the utility system, the voltages are generally the subject.

23.5.4 Power System Quantities under Nonsinusoidal Conditions

Traditional power system quantities such as rms, power (reactive, active, apparent), power factor, and phase sequences are defined for the fundamental frequency context in a pure sinusoidal condition. In the presence of harmonic distortion the power system no longer operates in a sinusoidal condition, and unfortunately many of the simplifications power engineers use for the fundamental frequency analysis do not apply. Therefore, these quantities must be redefined.

23.5.5 RMS Values of Voltage and Current

In a sinusoidal condition both the voltage and current waveforms contain only the fundamental frequency component, thus the rms values can be expressed simply as

$$V_{\text{rms}} = \frac{1}{\sqrt{2}} V_1 \qquad \text{and} \qquad I_{\text{rms}} = \frac{1}{\sqrt{2}} I_1. \tag{23-2}$$

where V_1 and I_1 are the amplitude of voltage and current waveforms, respectively. The subscript 1 denotes quantities in the fundamental frequency. In a nonsinusoidal condition a harmonically distorted waveform is made up of sinusoids of harmonic frequencies with different amplitudes as shown in Fig. 23-2. The rms values can of the waveforms are computed as the square root of the sum of rms squares of all individual components, that is,

$$V_{\text{rms}} = \sqrt{\sum_{h=1}^{h_{\max}} \left(\frac{1}{\sqrt{2}} V_h \right)^2} = \frac{1}{\sqrt{2}} \sqrt{V_1^2 + V_2^2 + V_3^2 + \cdots + V_{h_{\max}}^2}, \tag{23-3}$$

$$I_{\text{rms}} = \sqrt{\sum_{h=1}^{h_{\max}} \left(\frac{1}{\sqrt{2}} I_h \right)^2} = \frac{1}{\sqrt{2}} \sqrt{I_1^2 + I_2^2 + I_3^2 + \cdots + I_{h_{\max}}^2}, \tag{23-4}$$

where V_h and I_h are the amplitude of a waveform at the harmonic component h. In the sinusoidal condition, harmonic components of V_h and I_h are all zero, and only V_1 and I_1 remain. Equations (23-3) and (23-4) simplify to Eq. (23-2).

23.5.6 Active Power

The *active power*, P, is also commonly referred to as the average power, real power, or true power. It represents useful power expended by loads to perform real work, that is, to convert electric energy to other form of energy. Real work performed by an incandescent light bulb is to convert electric energy into light and heat. In electric power, real work is performed for the portion of the current that is in phase with the voltage. No real work will result, from the portion where the current is not in phase with the voltage. The active power is the rate at which energy is expended, dissipated or consumed by the load, and is measured in units of watts (W). P can be computed by averaging the product of the instantaneous voltage and current, that is,

$$P = \frac{1}{T}\int_0^T v(t)i(t)dt. \tag{23-5}$$

The above equation is valid for both sinusoidal and nonsinusoidal conditions. For sinusoidal condition, $I_{1\text{rms}}$ resolves to the familiar form,

$$P = \frac{V_1 I_1}{2}\cos\theta_1 = V_{1\text{rms}} I_{1\text{rms}} \cos\theta_1 = S\cos\theta_1, \tag{23-6}$$

where θ_1 is the phase angle between voltage and current at the fundamental frequency. Equation 23-6 indicates that the average active power is a function only of the fundamental frequency quantities. In the nonsinusoidal case, the computation of the active power must include contribution from all harmonic components, thus it is the sum of active power at each harmonic. Furthermore, because the voltage distortion is generally very low on power systems (less than 5%), Eq. (23-6) is a good approximation regardless of how distorted the current is. This approximation cannot be applied when computing the apparent and reactive power. These two quantities are greatly influenced by the distortion. The apparent power, S, is a measure of the potential impact of the load on the thermal capability of the system. It is proportional to the rms of the distorted current and it computation is straightforward, although slightly more complicated than the sinusoidal case. Also, many current probes can now directly report the true rms value of a distorted waveform.

23.5.7 Reactive Power

The *reactive power* is a type of power that does no real work and is generally associated with reactive elements (inductors and capacitors). For example, the inductance of a load such as a motor causes the load current to lag behind the voltage. Power appearing across the inductance sloshes back and forth between the inductance itself and the power system source producing no net work. For this reason it is called imaginary or reactive power since no power is dissipated or expended. It is expressed in units of volt-ampere-reactive or var. In the sinusoidal case, the reactive power is simply defined as

$$Q = S\sin\theta_1 = \frac{V_1 I_1}{2}\sin\theta_1 = V_{1\text{rms}} I_{1\text{rms}}\sin\theta_1 \tag{23-7}$$

which is the portion of power in quadrature with the active power shown in Eq. (23-6). Figure 23-34 summarizes the relationship between P, Q, and S in sinusoidal condition.

There is some disagreement among harmonics analysts on how to define Q in the presence of harmonic distortion. If it were not for the fact that many utilities measure Q and compute demand billing from the power factor computed by Q, it might be a moot point. It is more important to determine P and S; P defines how much active power is being consumed while S defines the capacity of the power system required to deliver P. Q is not actually very useful by itself. However, Q_1 the traditional reactive power component at fundamental frequency, may be used to size shunt capacitors.

The reactive power, when distortion is present, has another interesting peculiarity. In fact, it may not be appropriate to call it reactive power. The concept of var flow in the power system is deeply ingrained in the minds of most power engineers. What many do not realize is that this concept is valid only in the sinusoidal steady state. When distortion is present, the component of S that remains after P is taken out, is not conserved—that is, it does not sum to zero at a node. Power quantities are presumed to flow around the system in a conservative manner.

This does not imply that P is not conserved or that current is not conserved because the conservation of energy and Kirchoff's current laws are still applicable for any waveform. The reactive components actually sum in quadrature (square root of the sum of the squares). This has prompted some analysts to propose that Q be used to denote the reactive components that are conserved and introduce

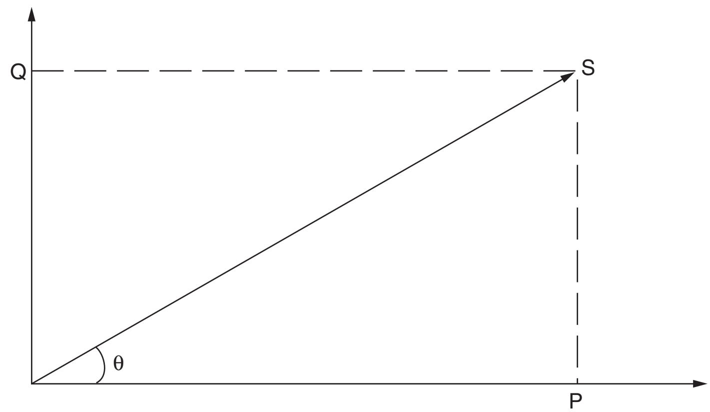


FIGURE 23-34 Relationship between P, Q, and S in sinusoidal condition.

a new quantity for the components that are not. Many call this quantity *D*, for *distortion power* or, simply, *distortion voltamperes*. It has units of voltamperes, but it may not be strictly appropriate to refer to this quantity as *power*, because it does not flow through the system as power is assumed to do. In this concept, *Q* consists of the sum of the traditional reactive power values at each frequency. *D* represents all cross products of voltage and current at different frequencies, which yield no average power. *P*, *Q* , *D*, and *S* are related as follows, using the definitions for *S* and *P* above as a starting point:

$$S = \sqrt{P^2 + Q^2 + D^2}$$
$$Q = \sum_k V_k I_k \sin \theta_k. \tag{23-8}$$

Therefore, *D* can be determined after *S*, *P*, and *Q* by

$$D = \sqrt{S^2 - P^2 - Q^2}. \tag{23-9}$$

Some prefer to use a three-dimensional vector chart to demonstrate the relationships of the components as shown in Fig. 23-35. *P* and *Q* contribute the traditional sinusoidal components to *S* while *D* represents the additional contribution to the apparent power by the harmonics.

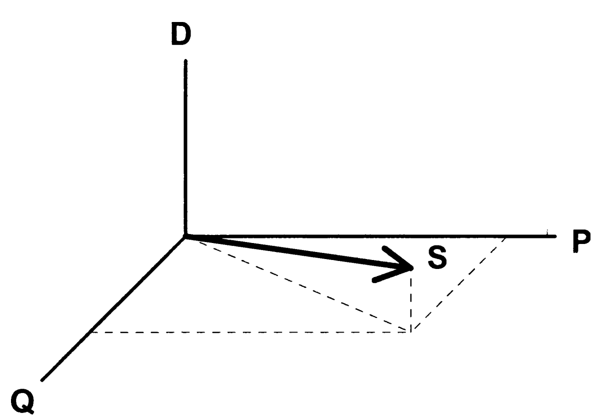


FIGURE 23-35 Relationship of components of the apparent power.

23.5.8 Power Factor

A *power factor* is a ratio of useful power to perform real work (active power) to the power supplied by a utility (apparent power), that is,

$$PF = \frac{P}{S} \tag{23-10}$$

In other words, the power factor ratio measures the percentage of power expended for its intended use. Power factor ranges from zero to unity. A load with power factor of 0.9 lagging denotes that the load can effectively expend 90% of the apparent power supplied (VA) and convert it to perform useful work (W). The term "lagging" denotes that the fundamental current lags behind the fundamental voltage by 25.84°.

In the sinusoidal case there is only one phase angle between the voltage and the current (since only the fundamental frequency is present), the power factor can be computed as the cosine of the phase angle and is commonly referred as the *displacement power factor*,

$$PF = \frac{P}{S} = \cos\theta. \tag{23-11}$$

In the nonsinusoidal case the power factor cannot be defined as the cosine of the phase angle as in Eq. (23-11). The power factor that takes into account contribution from all active power both fundamental and harmonic frequencies is known as the *true power factor*. The true power factor is simply the ratio of total active power for all frequencies to the apparent power delivered by the utility as shown in Eq. (23-10).

Power quality monitoring instruments now commonly report both displacement and true power factors. Many devices such as switch-mode power supplies and PWM adjustable-speed drives have a near-unity displacement power factor, but the true power factor may be 0.5 to 0.6. An ac-side capacitor will do little to improve the true power factor in this case because is Q_1 zero. In fact, if it results in resonance, the distortion may increase, causing the power factor to degrade. The true power factor indicates how large the power delivery system must be built to supply a given load. In this example, using only the displacement power factor would give a false sense of security that all is well.

The bottom line is that distortion results in additional current components flowing in the system that do not yield any net energy except that they cause losses in the power system elements they pass through. This requires the system to be built to a slightly larger capacity to deliver the power to the load than if no distortion were present.

23.5.9 Harmonic Phase Sequence

Power engineers have traditionally used symmetrical components to help describe 3-phase system behavior. The 3-phase system is transformed into three single-phase systems that are much simpler to analyze. The method of symmetrical components can be employed for analysis of the system's response to harmonic currents provided care is taken not to violate the fundamental assumptions of the method.

The method allows any unbalanced set of phase currents (or voltages) to be transformed into three balanced sets. The *positive sequence* set contains three sinusoids displaced 120° from each other, with the normal A-B-C phase rotation (e.g., 0°, –120°, 120°). The sinusoids of the *negative-sequence* set are also displaced 120°, but have opposite phase rotation (A-C-B, e.g., 0°, 120°, –120°). The sinusoids of the *zero sequence* are in phase with each other (e.g., 0, 0, 0).

In a perfect balanced 3-phase system, the harmonic phase sequence can be determined by multiplying the harmonic number h with the normal positive sequence phase rotation. For example, for the second harmonic, $h = 2$, produces 2 × (0°, –120°, –120°) or (0°, 120°, –120°) which is the negative sequence. For the third harmonic, $h = 3$, produces 3 × (0°, –120°, –120°) or (0°, 0°, 0°) which is the zero sequence. Phase sequence for all other harmonic orders can be determined in the same fashion. Since a distorted waveform in power systems contains only odd harmonic components (see Sec. 23.5.1), only odd harmonic phase sequence rotations are summarized below:

- Harmonics of order $h = 1, 7, 13, \ldots$ are purely positive sequence.
- Harmonics of order $h = 5, 11, 17, \ldots$ are purely negative sequence.
- Triplens ($h = 3, 9, 15, \ldots$) are purely zero sequence.

23.5.10 Triplen Harmonics

Triplen harmonics are the odd multiples of the third harmonic ($h = 3, 9, 15, 21, \ldots$). They deserve special consideration because the system response is often considerably different for triplens than for the rest of the harmonics. Triplens become an important issue for grounded-wye systems with current flowing on the neutral. Two typical problems are overloading the neutral and telephone interference. One also hears occasionally of devices that misoperate because the line-to-neutral voltage is badly distorted by the triplen harmonic voltage drop in the neutral conductor.

For the system with perfectly balanced single-phase loads illustrated in Fig. 23-36, an assumption is made that fundamental and third harmonic components are present. Summing the currents at node *N*, the fundamental current components in the neutral are found to be zero, but the third harmonic components are three times the phase currents because they naturally coincide in phase and time.

23.5.11 Triplen Harmonics in Transformers

Transformer winding connections have a significant impact on the flow of triplen harmonic currents from single-phase nonlinear loads. Two cases are shown in Fig. 23-37. In the wye-delta transformer (top), the triplen harmonic currents are shown entering the wye side. Since they are in phase, they add in the neutral. The delta winding provides ampere-turn balance so that they can flow, but they remain trapped in the delta and do not show up in the line currents on the delta side. When the currents are balanced, the triplen harmonic currents behave exactly as zero-sequence currents, which is precisely what they are. This type of transformer connection is the most common employed in utility

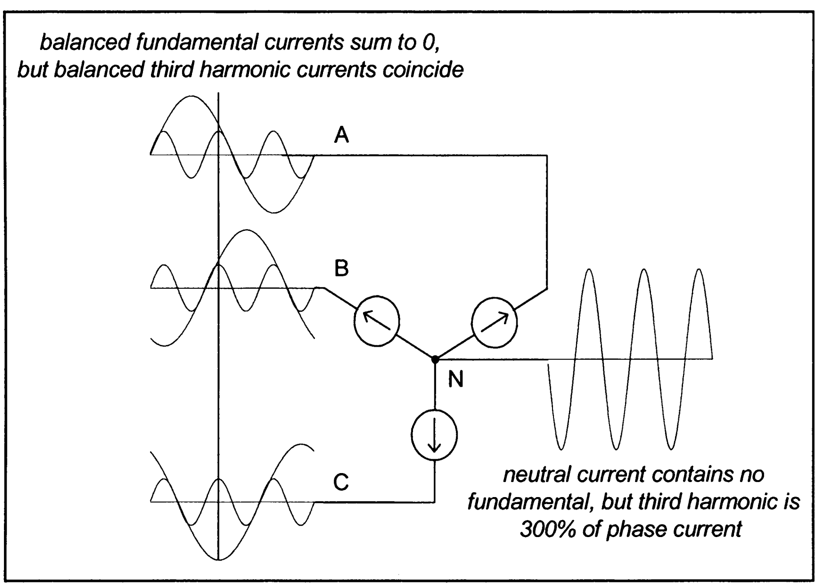


FIGURE 23-36 High neutral currents in circuits serving single-phase nonlinear loads.

distribution substations with the delta winding connected to the transmission feed.

Using grounded-wye windings on both sides of the transformer (bottom) allows balanced triplens to flow from the low voltage system to the high voltage system unimpeded. They will be present in equal proportion on both sides. Many loads in the United States are served in this fashion.

Some important implications of this related to power quality analysis are

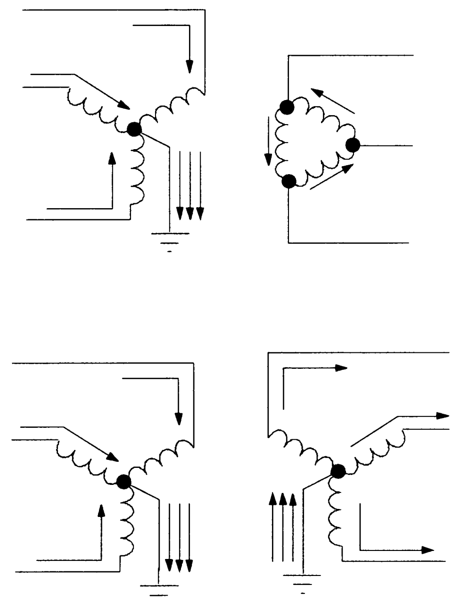


FIGURE 23-37 Flow of third harmonic current in 3-phase transformers.

1. Transformers, particularly the neutral connections, are susceptible to overheating when serving single phase loads on the wye side that have high third harmonic content.
2. Measuring the current on the delta side of a transformer will not show the triplens and, therefore, not give a true idea of the heating the transformer is being subjected to.

   The flow of triplen harmonic currents can be interrupted by the appropriate isolation transformer connection.
3. Removing the neutral connection in one or both wye windings, blocks the flow of triplen harmonic current. There is no place for ampere-turn balance. Likewise, a delta winding blocks the flow from the line. One should note that three-legged core transformers behave as if they have a "phantom" delta tertiary winding. Therefore, a wye-wye with only one neutral point grounded will still be able to conduct the triplen harmonics from that side.

These rules about triplen harmonic current flow in transformers apply only to balanced loading conditions. When the phases are not balanced, currents of normal triplen harmonic frequencies may very well show up where they are not expected. The normal mode for triplen harmonics is to be zero sequence. During imbalances, triplen harmonics may have positive or negative sequence components too.

One notable case of this is a 3-phase arc furnace. The furnace is nearly always fed by a delta-delta connected transformer to block the flow of the zero sequence currents, as shown in Fig. 23-8. Thinking that third harmonics are synonymous with zero sequence, many engineers are surprised to find substantial third harmonic current present in large magnitudes in the line current. However, during scrap meltdown, the furnace will frequently operate in an unbalanced mode with only two electrodes carrying current. Large third harmonic currents can then freely circulate in these two phases just as a single-phase circuit. However, they are not zero sequence currents. The third harmonic currents are equal amounts of positive and negative sequence currents. But to the extent that the system is mostly balanced, triplens mostly behave in the manner described.

23.5.12 Total Harmonic Distortion

The total harmonic distortion (THD) is a measure of the *effective value* of the harmonic components of a distorted waveform. That is, the potential heating value of the harmonics relative to the fundamental. This index can be calculated for either voltage or current:

$$\mathrm{THD} = \frac{\sqrt{\sum_{h>1}^{h_{\max}} M_h^2}}{M_1} \tag{23-12}$$

where M_h is the rms value of harmonic component h of the quantity M. The rms value of a distorted waveform is the square root of the sum of the squares as shown in Eq. (23-3) and (23-4). THD is related to the rms value of the waveform as follows:

$$\text{RMS} = \sqrt{\sum_{h=1}^{h_{max}} M_h^2} = M_1 \cdot \sqrt{1 + THD^2} \tag{23-13}$$

THD is a very useful quantity for many applications, but its limitations must be realized. It can provide a good idea of how much extra heat will be realized when a distorted voltage is applied across a resistive load. Likewise, it can give an indication of the addition losses caused by the current flowing through a conductor. However, it is not a good indicator of the voltage stress within a capacitor because that is related to the peak value of the voltage wave form, not its heating value. The THD index is most often used to describe voltage harmonic distortion. Harmonic voltages are almost always referenced to the fundamental value of the waveform at the time of the sample. Because voltage varies only a few percent, the voltage THD is nearly always a meaningful number.

23.5.13 Total Demand Distortion

Current distortion levels can be characterized by a THD value, as described above, but this can often be misleading. A small current may have a high THD but not be a significant threat to the system For example, many adjustable speed drives will exhibit high THD values for the input current when they are operating at very light loads. This is not necessarily a significant concern because the magnitude of harmonic current is low, even though its relative current distortion is high.

Some analysts have attempted to avoid this difficulty by referring THD to the fundamental of the peak demand load current rather than the fundamental of the present sample. This is called *total demand distortion* (*TDD*), and serves as the basis for the guidelines in IEEE STD 519-1992. It is defined as follows:

$$\text{TDD} = \frac{\sqrt{\sum_{h=2}^{h_{max}} I_h^2}}{I_L} \tag{23-14}$$

where I_L is the peak, or maximum demand load current at the fundamental frequency component measured at the point of common coupling (PCC). There are two ways to measure I_L. With a load already in the system, it can be calculated as the average of the maximum demand current for the preceding 12 months. The calculation can simply be done by averaging the 12-month peak demand readings. For a new facility, I_L has to be estimated based on the predicted load profiles.

23.5.14 System Response Characteristics

In analyzing harmonic problems, the response of the power system is equally as important as the sources of harmonics. In fact, power systems are quite tolerant of the currents injected by harmonic-producing loads unless there is some adverse interaction with the impedance of the system. Identifying the sources is only half the job of harmonic analysis. The response of the power system at each harmonic frequency determines the true impact of the nonlinear load on harmonic voltage distortion.

There are three primary variables affecting the system response characteristics, that is, the system impedance, the presence of capacitor bank, and the amount of resistive loads in the system.

23.5.15 System Impedance

At the fundamental frequency, power systems are primarily inductive, and the equivalent impedance is sometimes called simply the short-circuit reactance. Capacitive effects are frequently neglected on utility distribution systems and industrial power systems. One of most frequently-used quantities in the analysis of harmonics on power systems is the short-circuit impedance to the point on a network at which a capacitor is located. If not directly available, it can be computed from short-circuit study results that give either the short-circuit *MVA* or the short-circuit current as follows:

$$\begin{aligned} Z_{SC} &= R_{SC} + jX_{SC} \\ &= \frac{kV^2}{MVA_{SC}} = \frac{kV}{\sqrt{3}I_{SC}} \end{aligned} \tag{23-15}$$

where Z_{SC} = Short-circuit impedance
R_{SC} = Short-circuit resistance
X_{SC} = Short-circuit reactance
KV = Phase-to-phase voltage, kV
MVA_{SC} = 3-phase short-circuit, MVA
I_{SC} = Short-circuit current, A

Z_{SC} is a phasor quantity, consisting of both resistance and reactance. However, if the short-circuit data contains no phase information, one is usually constrained to assuming that the impedance is purely reactive. This is a reasonably good assumption for industrial power systems for buses close to the mains and for most utility systems. When this is not the case, an effort should be made to determine a more realistic resistance value because that will affect the results once capacitors are considered.

The inductive reactance portion of the impedance changes linearly with frequency. One common error made by novices in harmonic analysis is to forget to adjust the reactance for frequency. The reactance at the *h*-th harmonic is determined from the fundamental-impedance reactance, X_1, by

$$X_h = hX_1 \tag{23-16}$$

In most power systems, one can generally assume that the resistance does not change significantly when studying the effects of harmonics less than the ninth. For lines and cables, the resistance varies approximately by the square root of the frequency once skin effect becomes significant in the conductor at a higher frequency. The exception to this rule is with some transformers. Because of stray eddy current losses, the apparent resistance of larger transformers may vary almost proportionately with the frequency. This can have a very beneficial effect on damping of resonance as shown later. In smaller transformers, less than 100 kVA, the resistance of the winding is often so large relative to the other impedances that it swamps out the stray eddy current effects and there is little change in the total apparent resistance until the frequency reaches about 500 Hz. Of course, these smaller transformers may have an *X*/*R* ratio of 1.0 to 2.0 at fundamental frequency while large substation transformers might typically be 20 to 30. Therefore, if the bus that is being studied is dominated by transformer impedance rather than line impedance, the system impedance model should be considered more carefully. Neglecting the resistance will generally give a conservatively high prediction of the harmonic distortion.

At utilization voltages, such as industrial power systems, the equivalent system reactance is often dominated by the service transformer impedance. A good approximation for X_{SC} may be based on the impedance of the service entrance transformer only

$$X_{SC} \approx X_{tx} \tag{23-17}$$

While not precise, this is generally at least 90% of the total impedance and is commonly more. This is usually sufficient to evaluate whether or not there will be a significant harmonic resonance problem. Transformer impedance in ohms can be determined from the percent impedance, Z_{tx}, found on the nameplate by

$$X_{tx} = \left(\frac{kV^2}{MVA_{3\phi}}\right) \times Z_{tx}(\%) \tag{23-18}$$

where $MVA_{3\phi}$ is the kVA rating of the transformer. This assumes that the impedance is predominantly reactive. For example for a 1500 kVA, 6% transformer, the equivalent impedance on the 480 V side is

$$X_{tx} = \left(\frac{kV^2}{MVA_{3\phi}}\right) \times Z_{tx}(\%) = \left(\frac{0.480^2}{1.5}\right) \times 0.06 = 0.0092\Omega$$

23.5.16 Capacitor Impedance

Shunt capacitors, either at the customer location for power factor correction, or on the distribution system for voltage control, dramatically alter the system impedance variation with frequency. Capacitors do not create harmonics, but severe harmonic distortion can sometimes be attributed to their presence. While the reactance of inductive components increases proportionately to frequency, capacitive reactance, X_c, decreases proportionately:

$$X_C = \frac{1}{2\pi fC} \tag{23-19}$$

where C is the capacitance in farads. This quantity is seldom readily available for power capacitors, which are rated in terms of kvar or Mvar at a given voltage. The equivalent line-to-neutral capacitive reactance at fundamental frequency for a capacitor bank can be determined by

$$X_c = \frac{kV^2}{\text{M}var} \tag{23-20}$$

For 3-phase banks, use phase-to-phase voltage and the 3-phase reactive power rating. For single-phase units, use the can voltage rating and the reactive power rating. For example, for a 3-phase, 1200 kvar, 13.8-kV capacitor bank, the positive-sequence reactance in ohms would be

$$X_c = \frac{kV^2}{\text{M}var} = \frac{13.8^2}{1.2} = 158.7\Omega$$

23.5.17 Parallel and Series Resonance

All circuits containing both capacitance and inductance have one or more natural resonant frequencies. When one of these frequencies corresponds to an exciting frequency being produced by nonlinear loads, harmonic resonance can occur. Voltage and current will be dominated by the resonant frequency and can be highly distorted. Thus, the response of the power system at each harmonic frequency determines the true impact of the nonlinear load on harmonic voltage distortion.

Resonance can cause nuisance tripping of sensitive electronic loads and high harmonic currents in feeder capacitor banks. In severe cases, capacitors produce audible noise, and they sometimes bulge. To better understand resonance, consider the simple parallel and series cases shown in the one-line diagrams of Fig. 23-38. Parallel resonance occurs when the power system presents a parallel

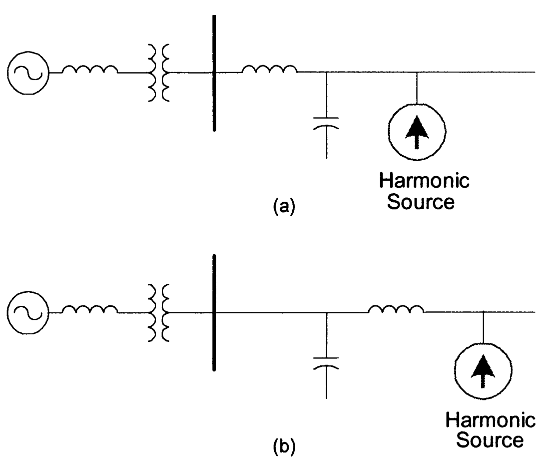


FIGURE 23-38 Examples of (*a*) parallel and (*b*) series resonance.

combination of power system inductance and power factor correction capacitors at the nonlinear load. The product of harmonic impedance and injection current produces high harmonic voltages. Series resonance occurs when the system inductance and capacitors are in series, or nearly in series, with respect to the nonlinear load point. For parallel resonance, the highest voltage distortion is at the nonlinear load. However, for series resonance, the highest voltage distortion is at a remote point, perhaps miles away or on an adjacent feeder served by the same substation transformer. Actual feeders can have five or ten shunt capacitors each, so many parallel and series paths exist, making computer simulations necessary to predict distortion levels throughout the feeder.

In the simplest parallel resonant cases, such as an industrial facility where the system impedance is dominated by the service transformer, shunt capacitors are located inside the facility, and distances are small. In these cases, the simple parallel scenario shown in Fig. 23-38*a* often applies.

23.5.18 Effects of Resistance and Resistive Load

Determining that the resonant harmonic aligns with a common harmonic source is not always cause for alarm. The damping provided by resistance in the system is often sufficient to prevent catastrophic voltages and currents. Figure 23-39 shows the parallel resonant circuit impedance characteristic for various amounts of resistive load in parallel with the capacitance. As little as 10% resistive loading can have a significant beneficial impact on peak impedance. Likewise, if there is a significant length of lines or cables between the capacitor bus and the nearest upline transformer, the resonance will be suppressed. Lines and cables can add a significant amount of the resistance to the equivalent circuit.

Loads and line resistances are the reasons why catastrophic harmonic problems from capacitors on utility distribution feeders are seldom seen. That is not to say that there will not be any harmonic problems due to resonance, but that the problems will generally not cause physical damage to the electrical system components. The most troublesome resonant conditions occur when capacitors are installed on substations buses, either utility substations or in industrial facilities. In these cases, where the transformer dominates the system impedance and has a high X/R ratio, the relative resistance is low and the corresponding parallel resonant impedance peak is very sharp and high. This is a common cause of capacitor failure, transformer failure, or the failure of load equipment.

It is a misconception that resistive loads damp harmonics because in the absence of resonance, loads of any kind will have little impact on the harmonic currents and resulting voltage distortion. Most of the current will flow back into the power source. However, it is very appropriate to say that resistive loads will damp *resonance*, which will lead to a significant reduction in the harmonic distortion.

23.5.19 Harmonic Impacts

Harmonics have a number of undesirable effects on power system components and loads. These fall into two basic categories: short-term and long-term. Short-term effects are usually the most noticeable and are related to excessive voltage distortion. On the other hand, long-term effects often go

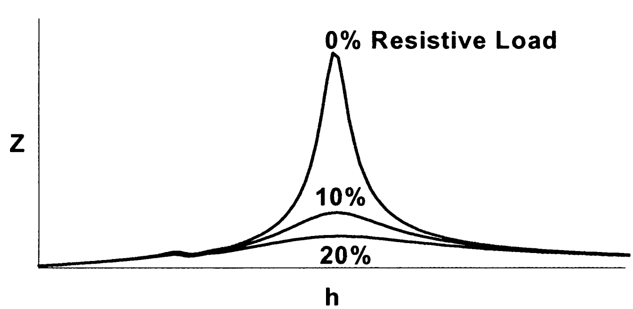


FIGURE 23-39 Effect of resistive loads on parallel resonance.

undetected and are usually related to increased resistive losses or voltage stresses. Short-term effects can cause nuisance tripping of sensitive loads. Some computer-controlled loads are sensitive to voltage distortion. For example, one documented case showed that a voltage distortion of 5.5% regularly shut down computerized lathes at a large pipe company heat treatment operation. While voltage distortions of 5% are not usually a problem, voltage distortions above 10% will almost always cause significant nuisance tripping or transformer overheating.

Harmonics can degrade meter accuracy. This is especially true with common single-phase induction-disk meters. In general, the meter spins 1% to 2% faster when a customer produces harmonic power. However, the greater issue in metering is the question of how active power, and especially reactive power, should be defined and measured when distortion is present. Debate on these definitions continues today.

Blown capacitor fuses and failed capacitor cans are also attributed to harmonics. Harmonic voltages produce excessive harmonic currents in capacitors because of the inverse relationship between capacitor impedance and frequency. Voltage distortions of 5% and 10% can easily increase rms currents by 10% to 50%. Capacitors may also fail because of overvoltage stress on dielectrics. A 10% harmonic voltage for any harmonic above the third increases the peak voltage by approximately 10% because the peak of the harmonic usually coincides, or nearly coincides, with the peak of the fundamental voltage.

Harmonics can also cause transformer overheating. This usually occurs when a dedicated transformer serves only one large nonlinear load. In such a situation, the transformer must be derated accordingly. Derating to 0.80 of nameplate kVA is common.

Overloaded neutrals appear to be the most common problems in commercial buildings. In a 3-phase, four-wire system, the sum of the 3-phase currents returns through the neutral conductor. Positive and negative sequence components add to zero at the neutral point, but zero sequence components are additive at the neutral.

23.5.20 Control of Harmonics

There are two common causes of harmonic problems are

- Nonlinear loads injecting excessive harmonic currents
- The interaction between harmonic currents and the system frequency response

When harmonics become a problem, commonly-employed solutions are

- Limit harmonic current injection from nonlinear loads. Transformer connections can be employed to reduce harmonics in a 3-phase system by using parallel delta-delta and wye-delta transformers to yield net 12-pulse operation, or delta connected transformers to block triplen harmonics.
- Modify system frequency response to avoid adverse interaction with harmonic currents. This can be done by feeder sectionalizing, adding or removing capacitor banks, changing the size of the capacitor banks, adding shunt filters, or adding reactors to detune the system away from harmful resonances.
- Filter harmonic currents at the load or on the system with shunt filters, or try to block the harmonic currents produced by loads. There are a number of devices to do this. Their selection is largely dependent on the nature of the problems encountered. Solutions can be as simple as an in-line reactor (i.e., a choke) as in PWM-based adjustable speed drive applications, or as complex as an active filter.

23.6 ELECTRICAL POWER RELIABILITY AND RECENT BULK POWER OUTAGES

23.6.1 Electric Power Distribution Reliability—General

The term *reliability* in the utility context usually refers to the amount of time end users are totally without power for an extended period of time (i.e., a sustained interruption). Definitions of what constitutes a sustained interruption vary among utilities in the range of 1 to 5 min. This is what many utilities refer to as an "outage." Current power quality standards efforts are leaning toward calling any interruption of power for longer than 1 min, a sustained interruption. In any case, reliability is affected by the permanent faults on the system that must be repaired before service can be restored.

23.6.2 Electric Power Distribution Reliability Indices

Most commonly used reliability indices for utility distribution systems are defined as follows:

- SAIFI: System Average Interruption Frequency Index

  SAIFI represents the average interruption frequency experienced by customers served in the system over a given period of time. It is computed as follows:

$$\text{SAIFI} = \frac{\text{(no. of customer interrupted)(no. of interruption)}}{\text{total no. of customers}}$$

- SAIDI: System Average Interruption Duration Index

  SAIDI represents the average interruption duration experienced by customers in the system over a given period of time.

$$\text{SAIDI} = \frac{\sum\text{(no. of customer affected)(duration of outage)}}{\text{total no. of customers}}$$

- CAIFI: Customer Average Interruption Frequency Index

  CAIFI represents average interruption frequency for affected customers. Customers not experiencing interruption are not included in the calculation.

$$\text{CAIFI} = \frac{\text{total no. of customer interruptions}}{\text{total no. of customers affected}}$$

- CAIDI: Customer Average Interruption Duration Index

  CAIDI represents the average interruption duration for customers experiencing interruptions. In other words, this is the average restoration time for affected customers.

$$\text{CAIDI} = \frac{\sum\text{(customer interruption durations)}}{\text{total no. of customers interruptions}}$$

- ASAI: Average System Availability Index

  ASAI represents the average system availability over a given observation period, which is usually a year (or 8760 hours). The index is given in percent.

$$\text{ASAI} = \frac{\sum\text{ customer hours service availability}}{\text{customer hours service demand}}$$

23.6.3 Major Bulk Electric Power Outages

Since the electric power industry was born in the early twentieth century, there have been several notable major bulk power outages. Most common causes of these outages are errors in protective device system design, overgrown vegetation, loss of system awareness due to failure of alarm systems, and a combination of unexpected events, whether they are natural and man made. Summary of these bulk power outages are complied from various sources and presented in the next paragraphs.

23.6.4 Great Northeast Blackout of 1965

The 1965 power outage started on November 9 at about 5:15 P.M. in Ontario, Canada. It cascaded down through the power system to the majority of New York, Connecticut, Massachusetts, Rhode Island, and some portions of northern Pennsylvania, and New Jersey. They were about 30 million customers out of service for up to 13 h. The power outage left 20 GW of load demand unserved. The outage was triggered by a backup protectively relay in opening one of five 230-kW lines delivering power from the Adam Beck Station No. 2 to the Toronto load area. System operators were not aware that the backup relay was set to take the line out of service when the line loading exceeded 375 MW. This relay setting was below the unusually high line loadings of recent months. Higher than normal line loadings was imposed due to higher than normal import from the United States to cover nearby Lakeview power plant (west of Toronto) outage. Upon opening the 230-kV line, the remaining four 230-kV lines were also tripped out successively within $2^1/_2$ s. Subsequently, two key east–west 345-kV lines between Rochester and Syracuse tripped out due to line instability. Several lower voltage lines tripped open along with 5 of 11 generation units at the St. Lawrence (Massena) Station. Losses of major transmission lines caused 10 generators at Adam Beck Station to shut down due to low governor oil pressure. By 5:30 P.M., the majority of northeast was without power. The service was, however, restored by 4:44 A.M. the next day in Manhattan [18].

23.6.5 New York Blackout of 1977

The event started on July 13 at about 8:37 P.M., when a lightning stroke caused a phase B to ground fault on both of a double-circuit 345-kV transmission line between Buchanan South and Millwood West Substations [18,19,21]. The tripping of circuit breakers at Buchanan South Circuit rings isolated Indian Point No. 3 generating unit without a transmission path to any load. The plant tripped of line and shut down causing a generation loss of 883 MW. A coordination error in the protective system played a critical role in the subsequent chain of events in which a transfer trip signal to Ladentown was initiated to open the 345-kV line from Buchanan South to Ladentown. A subsequent lighting stroke also caused a trip out of two more 345-kV lines between Sprain Brook and Buchanan North, and Sprain Brook and Millwood West. The later was restored to service in about 2 s. However, Sprain Brook to Buchanan North 345-kV was out of service. Losses of key transmission lines eventually forced the electrical system to separate and collapse. The power outage affected 9 million people. However, it was limited to New York City alone. Unlike the 1965 blackout, the 1977 event was marred by violence and looting [20].

Timeline of key events in the total collapse of the ConEd system [19,21]:

- At 8:37:17 P.M., July 13, 1997, two 345-kV lines connecting Buchanan South to Millwood West were each subjected to a phase B fault to ground as a result of a severe lightning stroke.
- The tripping of circuit breakers at the Buchanan South ring bus, isolated the Indian Point No. 3 generating unit from any load, and the unit tripped for a loss of 883 MW.
- Loss of the ring bus isolated the 345-kV tie to Ladentown, which had been importing 427 MW, with a total loss now of 1310 MW.
- At 8:55:53 P.M., about $18^1/_2$ minutes after the first incident, a severe lightning stroke caused the trip-out of two 345-kV lines, which connect Sprain Brook to Buchanan North, and Sprain Brook to Millwood West. These two 345-kV lines share common towers between Millwood West and Sprain Brook. One line (Millwood West to Sprain Brook) was restored to service in about 2 s. The failure of the other line to reclose isolated the last ConEd interconnection to the northwest.
- The resulting surge of power from the northwest, caused the trip-out of the line between Pleasant Valley and Millwood West (a bent contact on one of the relays at Millwood West caused the improper action).
- At 9:19:11 P.M., a 345-kV line, Leeds Substation to Pleasant Valley tripped as a result of a phase B fault to ground (fault probably caused by line sag to a tree because of the excessive overload imposed on the line).

- At 9:19:53 P.M., the 345-kV/138-kV transformer at Pleasant Valley tripped on overcurrent relay and left ConEd with three remaining interconnections.
- At 9:22:11 P.M., the Jamaica/Valley Stream tie was opened manually by the Long Island Lighting Co. system operator after obtaining the approval of the pool dispatcher.
- About 7 min later, the tap-changing mechanism failed on the Goethals phase-angle regulating transformer resulting in the trip of the Linden/Goethals tie to PJM, which was carrying 1150 MW to ConEd.
- The two remaining external 138-kV ties to ConEd tripped on overload isolating the ConEd system.

23.6.6 The Northwestern Blackout of July 1996

This outage occurred on July 2 at about 2:24 P.M. when a tree fault tripped a 345-kV taking power from Jim Bridger power plant in southwest Wyoming to southeast Idaho [17]. Protective devices detected the fault and de-energized the line. However, due to a protection coordination error, a parallel 345-kV line was also tripped. The loss of two 345-kV line severely limited power transfers from Jim Bridger plant causing generator protective devices to trip two 500-MW generators to maintain the system stability. With two generators out, frequency in the entire western interconnection began to decline forcing some customers out of service. This move was not successful and the system disintegrated into five islands. About 2 million customers were interrupted for up to several hours with about 11,850 MW of loss of load demand.

23.6.7 The Northwestern Blackout of August 1996

This blackout occurred on August 10 when Keeler-Allston 500-kV transmission line sagged into a grove of trees [17]. Prior to the disturbance, the Northwest area was importing about 2300 MW from Canada due to excellent hydroelectric conditions that lead to high electricity transfers. This and other conditions, that is, hot weather, maintenance outage of a transformer that connects a static var compensator to a 500-kV line in Portland, and/or failure to trim trees, lead to a cascading outage. A series tree fault disturbances finally broke the western interconnection area into four islands and interrupting services to 7.5 million customers for up to 9 h.

23.6.8 The Great Northeastern Power Blackout of 2003 [22, 23]

This outage on August 14, 2003 is by far the largest and most severe among all major outages. It affected 50 million customers for up to 2 days. The area affected was over 9266-mi [2] covering two Canadian provinces and eight northeastern U.S. states. The outage caused an estimated of \$4 to \$8 billion in lost economic activity.

The outage was preceded by a computer software abnormal operation, a series of generator tripping, and line outages. This series of events is considered as a precursor to the cascading outage. MISO's (Midwest Independent Service Operator) state estimator software solution did not converge and produced a solution with a high mismatch due to an outdated input data in the state estimator. Eastlake No. 5 generating unit tripped along with two other units (Conesville and Greenwood) causing a severe shortage in reactive power supply. Adequate reactive power supply is an important requirement high voltage long distance electric power transmission. At about 2:02 P.M., Stuart–Atlanta 345 kV in southwestern Ohio tripped due to contact with trees causing a short circuit to ground and locked out. This situation was exacerbated by lost of key alarm functions in FirstEnergy's (Ohio) control room. The controller also lost a series of other important computer functions. Unfortunately, FirstEnergy operators were unaware of computer failures thus they lost situational awareness of their system.

Precipitating events that lead to the cascading outage began around 3:00 P.M., when three key 345-kV transmission lines into northern Ohio from eastern Ohio tripped out, Harding–Chamberlain (3:05 P.M.), Hanna–Juniper (3:32 P.M.), and Start–South Canton (3:41 P.M.). They were all tripped out due to tree faults caused by overgrown vegetation. With these three 345-kV line out, power flowed

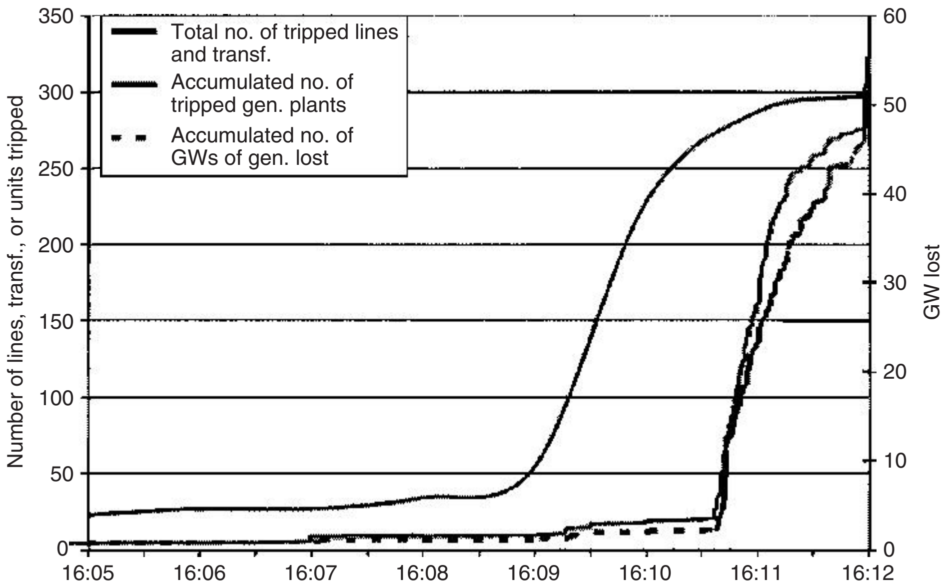


FIGURE 23-40 Accumulated line and generator trips during the cascade [23].

over through other remaining lines including the underlying 138-kV system. This redirection of flow caused severe overloading in 138-kV lines. As a result between 3:39 P.M. and 3:58 P.M., seven 138-kV lines tripped. At 3:59 P.M., West Akron bus tripped due to breaker failure. This event caused another five 138 kV-lines to trip. A few minutes later, between 4:00 P.M. and 4:08 P.M., another four 138-kV lines tripped. Losses of these transmission lines disconnected northern Ohio from Eastern Ohio. The last 345-kV line between Sammis and Star tripped at 4:06 P.M. The loss of this line left northern Ohio without any 345-kV path to eastern Ohio, and initiated a cascading blackout across the northeast U.S. and Canada. Within 7 min after the loss of 345-kV Sammis–Start line, more than 508 generating units at 265 power plants had been lost, and close to 300 lines and transformers tripped.

23.6.9 Power Quality Characteristics in the Great Northeastern Power Blackout of 2003

A major power quality characteristic of the blackout was sagging voltage when transmission lines experienced fault clearing operations (opening and reclosing) due to tree contacts.

An 8-cycle voltage sag was measured at an industrial site in Cleveland at about 3:45 P.M. This was when a series of 138-kV lines experience tree faults and attempted to isolate them (Fig. 23-41). This particular fault was detected and cleared promptly, but the voltage recovery at the site appears to be slow, suggesting that the system was now much weaker than previously. At least one more line was out of service.

Shortly after 4:00 P.M. there was another instantaneous voltage sag recorded (Fig. 23-42). The voltage drops abruptly and remains at the lower level. Phase unbalance develops, suggesting either the presence of a remote fault or that the system has become very weak due to the tripping of another line.

A short time later (accuracy of the time stamp is uncertain), the disturbance in the voltage shown in Fig. 23-43 was recorded in an office building in downtown Manhattan. This waveform is consistent with that of a power system that is going unstable.

Waveforms of this type can be observed in system dynamics simulations for buses in the weaker part of the system as it begins to move relative to a more distant part of the system that remains in

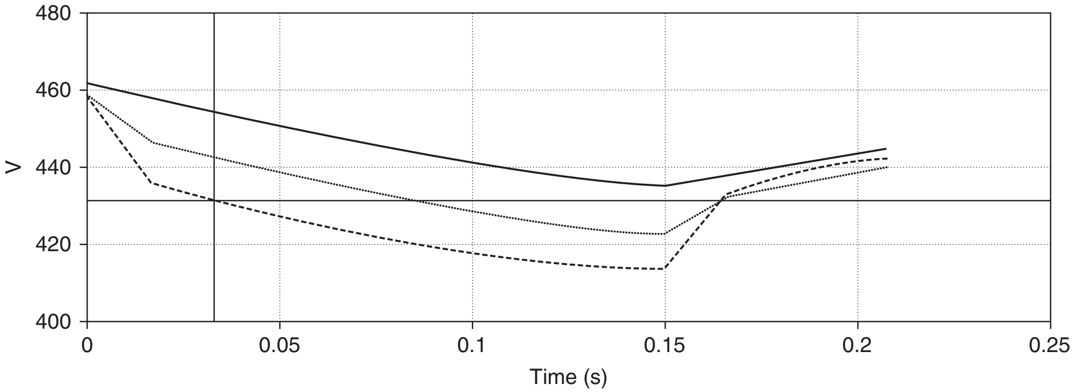


FIGURE 23-41 Instantaneous sag on phase C captured at service entrance of Cleveland industrial facility at around 3:45 P.M. (note time stamp is off by about 12 min) likely due to a fault on a transmission line. Voltage recovers slowly after fault is cleared, suggesting a weakened system. (*Courtesy of Electrotek Concepts and Dranetz-BMI*.)

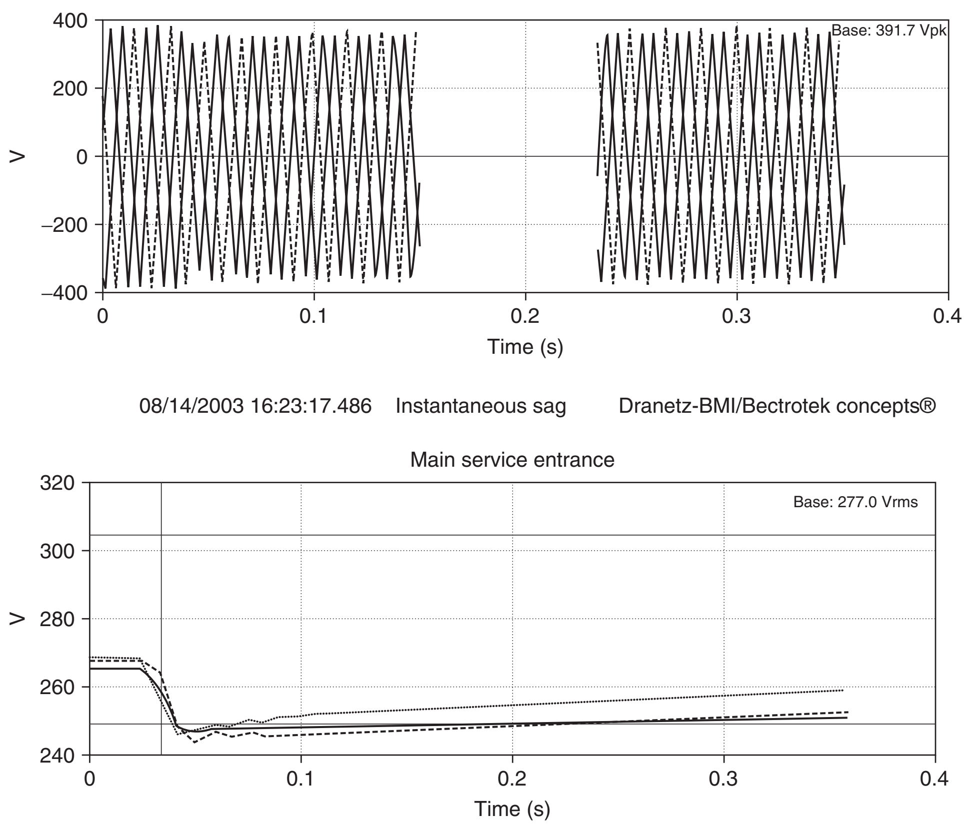


FIGURE 23-42 3-Phase instantaneous sag shortly after 4 P.M. on August 14. (*Courtesy of Electrotek Concepts and Dranetz-BMI*.)

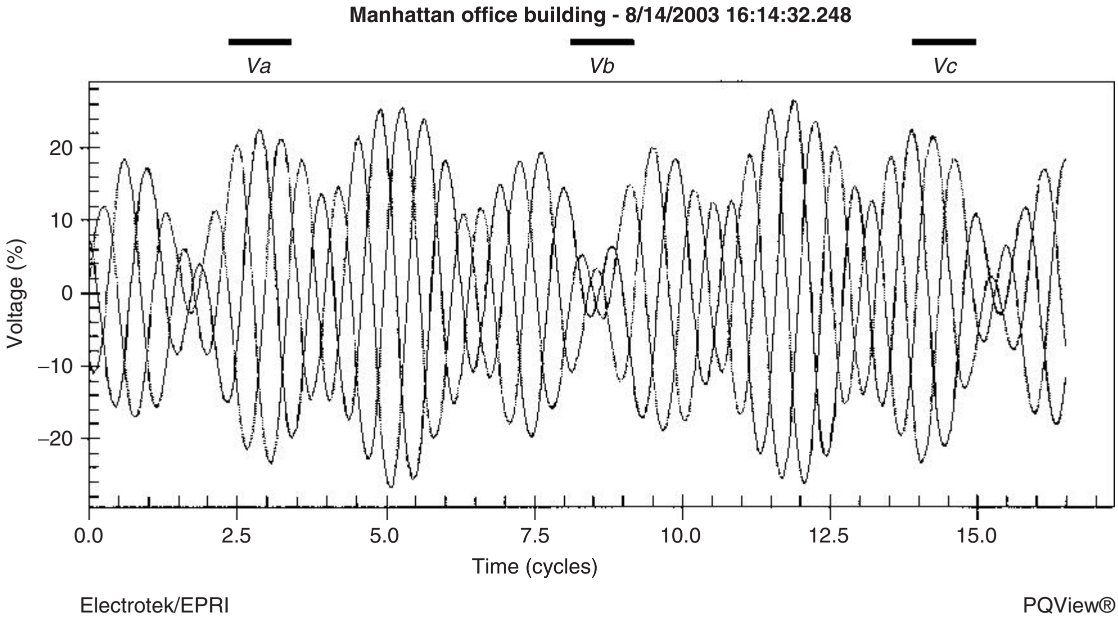


FIGURE 23-43 Waveform captured as the grid collapsed in New York City showing the instability of the system. (*Courtesy of Electrotek Concepts and Dranetz-BMI.*)

synchronism. A "beating" frequency develops as the two interconnected systems operate at different frequencies. In this case, it would appear that the system containing this power monitor had drifted by approximately 2 Hz from the rest of the system. Once this occurred, the power system supplying Manhattan detected the instability and immediately shut down the generators and separated from surrounding power systems.

Several of the neighboring systems successfully separated and remained stable throughout the blackout period despite briefly experiencing voltage waveforms like this while they were still interconnected with the part of the system that was collapsing.

Once the massive amount of load in the affected areas was lost, the entire eastern interconnection experienced a jump in frequency of approximately 0.2 Hz. This could be seen over a large geographic area (Fig. 23-44). After a few minutes, generator controls brought the average frequency back to 60 Hz and few energy users outside the affected area realized that anything had happened. While large in terms of system dynamics issues, this frequency change is inconsequential to most loads.

Figure 23-45 shows the complete rms voltage trend for the Manhattan site from the beginning of the blackout shortly after 4:00 P.M. on August 14 until the power was restored at 5:30 A.M. on August 15.

REFERENCES

1. TC77WG6 (Secretary) 110-R5, Draft Classification of Electromagnetic Environments, January 1991.
2. IEEE Std 1159-1995, Recommended Practice on Monitoring Electric Power.
3. IEC 50 (161) (1989), International Electrotechnical Vocabulary, Chapter 161: Electromagnetic Compatibility.
4. UIE-DWG-3-92-G, "Guide to Quality of Electrical Supply for Industrial Installations," Part 1: General Introduction to Electromagnetic Compatibility (EMC), Types of Disturbances and Relevant Standards. Advance UIE Edition. "Disturbances" Working Group GT 2.

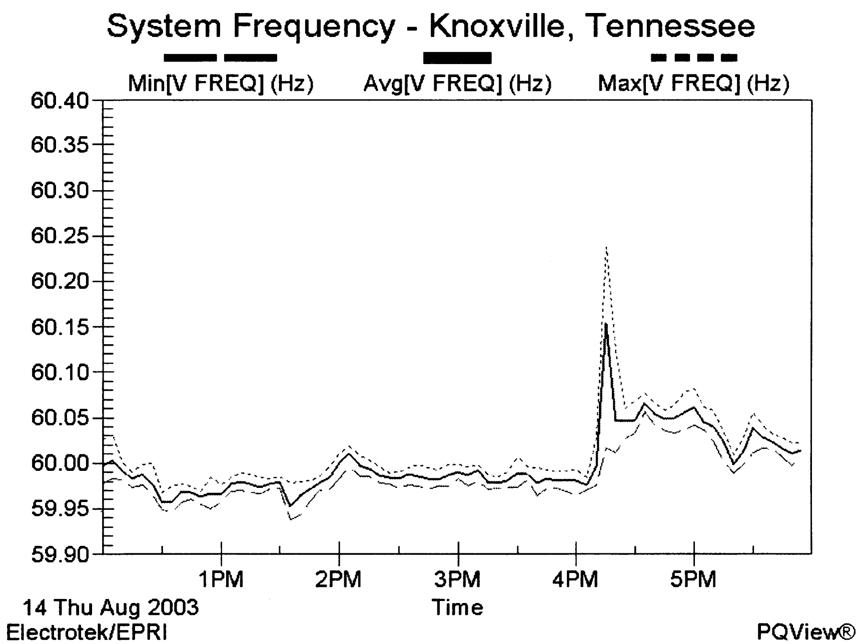


FIGURE 23-44 System frequency jump when the major systems separated was seen in locations as far away as Knoxville, TN. (*Courtesy of Electrotek Concepts and Dranetz-BMI.*)

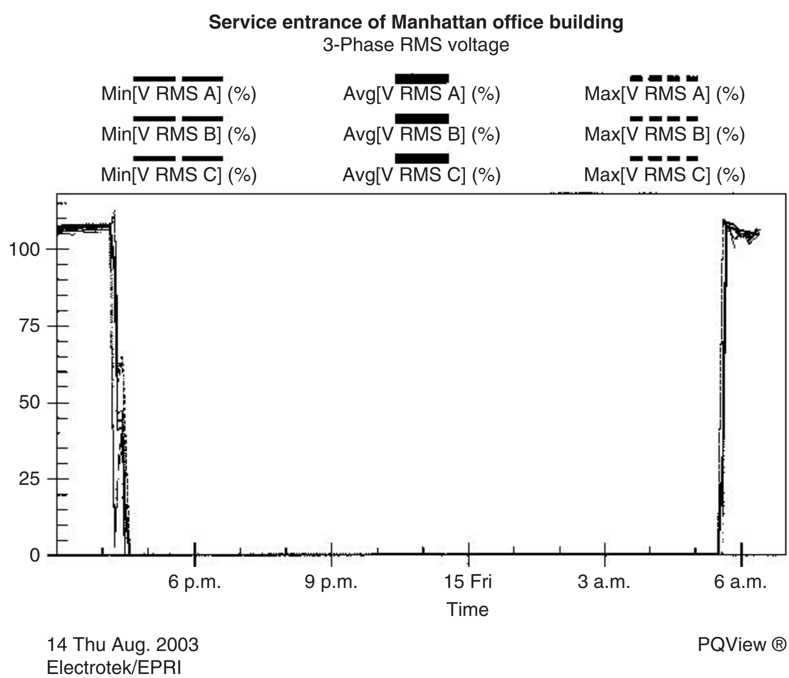


FIGURE 23-45 Rms voltage trend showing the min/max/avg of all three phases for the entire duration of the blackout at a Manhattan office building. (*Courtesy of Electrotek Concepts and Dranetz-BMI.*)

5. UIE-DWG-2-92-D, UIE Guide to Measurements of Voltage Dips and Short Interruptions Occurring in Industrial Installations.
6. IEC 61000-2-1(1990-05), "Description of the environment—Electromagnetic environment for low frequency conducted disturbances and signaling in public power supply systems," Electromagnetic Compatibility (EMC)—Part 2 Environment, Section 1, 1990.
7. ANSI/NFPA 70-1993, National Electrical Code.
8. IEEE Std 100-1992, IEEE Standard Dictionary of Electrical and Electronic Terms.
9. IEEE Std 446-1987, IEEE Recommended Practice for Emergency and Standby Power Systems for Industrial and Commercial Applications (IEEE Orange Book).
10. Information Technology Industry Council (ITI), 1250 Eye Street NW, Suite 200, Washington, DC, http://www.itic.org
11. IEC 61000-4-30 77A/356/CDV, Power Quality Measurement Methods.
12. IEC 61000-4-15, Flicker Meter—Functional and Design Specifications.
13. J. Lamoree, D. Mueller, P. Vinett, W. Jones, "Voltage Sag Analysis Case Studies," 1993 IEEE I&CPS Conference, St. Petersburg, FL.
14. R. C. Dugan, L. A. Ray, D. D. Sabin, et al., "Impact of Fast Tripping of Utility Breakers on Industrial Load Interruptions," Conference Record of the 1994 IEEE/IAS Annual Meeting, Vol III, Denver, October 1994, pp. 2326–2333.
15. IEEE Transformers Committee, "Secondary (Low-Side) Surges in Distribution Transformers," *Proceedings of the 1991 IEEE PES Transmission and Distribution Conference*, Dallas, September 1991, pp. 998–1008.
16. G. L. Goedde, R. C. Dugan, L. D. Rowe, "Full Scale Lightning Surge Tests of Distribution Transformers and Secondary Systems," *Proceedings of the 1991 IEEE PES Transmission and Distribution Conference,* Dallas, September, 1991, pp. 691–97.
17. Consortium for Electric Reliability Technology Solutions, "Review of Recent Reliability Issues and System Events," Grid of the Future White Paper, December, 1999.
18. Web site http://blackout.gmu.edu, accessed on February 24, 2005.
19. U.S. Department of Energy, The Con Edison Power Failure of July 13 and 14, 1977, Final Report, June 1978.
20. TIME, "Nigh of Terror," July 25, 1977, pp. 12–26.
21. Lesson Learned From the 1977 Blackout Case Study 1—Sequence of Events, www.blackout.gmu.edu/archive/pdf/lessons_learned_77.pdf, accessed on February 24, 2005.
22. North American Electric Reliability Council, Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn, July 13, 2004.
23. U.S.–Canada Power System Outage Task Force, Final Report on the August 14, 2003, Blackout in the United States and Canada: Causes and Recommendation, April 2004.

SECTION 24
GROUNDING SYSTEMS

David R. Stockin
Manager of Engineering, E&S Grounding Solutions

Michael A. Esparza
Principal & Director of Sales, E&S Grounding Solutions

Illustrations by **Gil Juarez,** RevDesign

CONTENTS

24.1 INTRODUCTION

In the last few decades, much has been ascertained about the interaction between the grounding electrode and the earth, which is a three-dimensional electrical circuit. Ultimately, it is the soil resistivity (and special variations thereof) that determines system design and performance. New technology has significantly reduced the resistance between grounding electrodes and the surrounding soil, which is a determining factor in the performance of small electrodes. There are a number of different grounding electrodes in use today. They are the standard driven rod, advanced driven rod, grounding plate, Ufer (concrete encased electrode), water pipes, and the electrolytic electrode.

The National Electric Code (NEC) divides grounding into two distinct areas: *equipment grounding* and *system grounding*. Equipment grounding is the process of connecting above-ground equipment to the earth. In other words, how to properly bond wires to equipment, routing them through conduits, circuit-breaker boxes, etc. System grounding is the process of intentionally making an electrical connection to the earth itself. This is the actual connection of metal to soil, and the minimum standards by which this connection is made. This process is often referred to as *earthing*.

The goal for this chapter is to provide a basic knowledge of system grounding and earthing in an easy-to-read and understandable manner. Above-ground wiring issues, except where needed is not discussed. The topics covered are system grounding, the benefits and features of the available grounding electrodes, and the ground potential rise (GPR) hazards of high current discharges. We also introduce the principles of proper soil testing, resistance-to-ground (RTG) testing, and meter selection.

Both equipment grounding and system grounding are becoming more essential as technology rapidly advances. Many of the latest and most advanced systems have stringent grounding requirements. Understanding the available electrical data through proper ground testing enables the electrical engineer to manage grounding systems that will meet specified grounding criteria.

Our goal is to provide the basic knowledge needed to understand and make the right choices when it comes to electrical grounding. Remember, "To protect what's above the ground you need to know what's in the ground."

24.2 SPHERE OF INFLUENCE

An important concept as to how efficiently grounding electrodes discharge electrons into the earth is called "the zone of influence," which is sometimes referred to as "the sphere of influence." The zone of influence (Fig. 24-1) is the volume of soil throughout which the electrical potential rises to more than a small percentage of the potential rise of the ground electrode, when that electrode discharges current into the soil. The greater the volume compared with the volume of the electrode, the more efficient the electrode. Elongated electrodes, such as ground rods, are the most efficient. The surface area of the electrode determines the ampacity of the device, but does not affect "the zone of influence." The greater the surface area of the electrode, the greater the contact with the soil and more electrical energy can be discharged per unit of time.

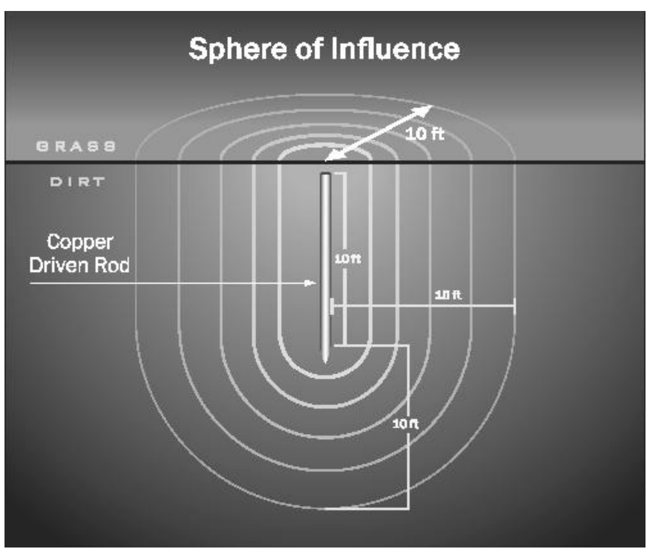


FIGURE 24-1 Sphere of influence of an earth electrode.

The formula for calculating the volume of soil is

$$V = 5\pi L^3/3 \tag{24-1}$$

where V = the volume of soil in the sphere of influence and L = the length or depth of the electrode.

A simpler version of the equation is used when the above formula is modified by rounding π(pi) down to 3 and cross canceling to get the formula:

$$V = 5L3 \tag{24-2}$$

Thus, a single 10-ft driven rod will utilize 5000 cubic ft of soil and a single 8-ft rod will utilize about half the soil at 2560 cubic ft.

24.3 GROUNDING ELECTRODES

Grounding is the process of electrically connecting any metallic object to the earth by the way of an earth electrode system. The *NEC* requires that the grounding electrodes be tested to ensure that they are under 25-Ω RTG (earth).

It is important to know, that aluminum electrodes are not allowed for use in grounding (Table 24-1).

24.3.1 Driven Rod

The standard driven rod or copper-clad rod (Fig. 24-2) consists of an 8- to 10-ft length of steel with a 5- to 10-mil coating of copper. This is by far the most common grounding device used in the field today. The driven rod has been in use since the earliest days of electricity with a history dating as far back as Benjamin Franklin.

TABLE 24-1 Earth-Electrode Comparison Chart

	Driven rod	Advanced driven rod	Grounding plate	Concrete encased electrode	Building foundation	Water pipe	Electrolytic electrode
Resistance-to-ground (RTG)	Poor	Average	Poor	Average	Above average	Poor	Excellent
Corrosion resistance	Poor	Good	Poor	Good*	Good*	Varies	High
Increase in RTG in cold weather	Highly affected	Slightly affected	Highly affected	Slightly affected	Highly affected	Varies w/ contact to earth	Minimally affected
Increase in RTG over time	RTG worsens	RTG typically unaffected	RTG increases	RTG typically unaffected	RTG typically unaffected	RTG typically unaffected	RTG improves
Electrode ampacity	Poor	Average	Average	Average*	Above average*	Poor	Excellent
Installation cost	Average	Excellent	Below average	Below average	Average	Average	Poor
Life expectancy	Poor 5–10 years	Average 15–20 years	Poor 5–10 years	average* 15–20 years	Above average* 20–30 years	Below average* 10–15 years	Excellent 30–50 years

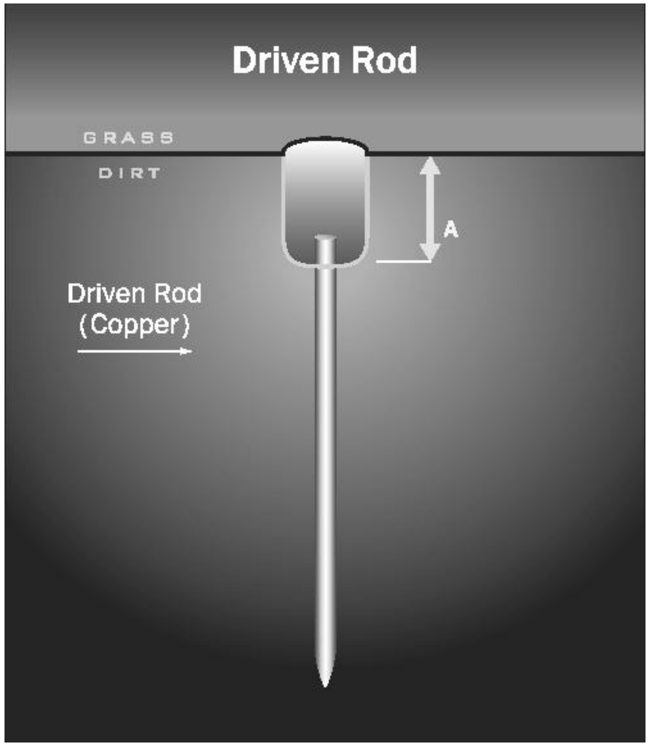


FIGURE 24-2 Copper-clad driven grounding rod.

Driven rods are relatively inexpensive to purchase, however, ease of installation is dependent upon the type of soil and terrain where the rod is to be installed. The steel used in the manufacture of a standard driven rod tends to be relatively soft. Mushrooming can occur on both the tip of the rod as it encounters rocks on its way down, and the end where force is being applied to drive the rod through the earth. Driving these rods can be extremely labor-intensive as rocky terrain creates problems as the tips of the rods continue to mushroom.

Often these rods will hit a rock and actually turn back around on themselves and pop back up a few feet away from the installation point. Because driven rods range in length from 8 to 10 ft, often a ladder is required to reach the top of the rod, which can become a safety issue. Many falls have occurred from personnel trying to literally "whack" these rods into the earth while hanging from a ladder many feet in the air.

The *NEC* requires that driven rods be a minimum of 8 ft in length and that 8 ft of length must be in direct contact with the soil. Typically, a shovel is used to dig down into the ground 18 in before a driven rod is installed. The most common rods used by commercial and industrial contractors are in 10-ft lengths. Many industrial specifications require this length as a minimum.

A common misconception is that the copper coating on a standard driven rod has been applied for electrical reasons. While copper is certainly a conductive material, its real purpose on the rod is to provide corrosion protection for the steel underneath. Many corrosion problems can occur because copper is not always the best choice in corrosion protection. It should be noted that galvanized driven rods have been developed to address the corrosion concerns that copper presents, and in many cases are a better choice for prolonging the life of the grounding rod and grounding systems. Generally speaking, galvanized rods are a better choice in all but high salt environments.

An additional drawback of the copper-clad driven rod is that copper and steel are two dissimilar metals. When an electrical current is imposed, electrolysis will occur. Additionally, the act of driving the rod into the soil can damage the copper cladding, allowing corrosive elements in the soil to attack the bared steel and further decrease the life expectancy of the rod. Environment, aging, temperature, and moisture also easily affect driven rods, giving them a typical life expectancy of 5 to 15 years in good soil conditions. Driven rods also have a very small surface area and that is not always conducive to good contact with the soil. This is especially true in rocky soil conditions where the rod will only make contact on the edges of the surrounding rock.

A good example of this is to imagine a driven rod surrounded by large marbles. Actual contact between the marbles and the driven rod will be very small. Because of this small surface contact with the surrounding soil, the RTG will increase, lowering the conductance, and limiting its ability to handle high-current faults.

24.3.2 Advanced Driven Rods

Advanced driven rods (Fig. 24-3) are specially engineered varieties of the standard driven rod with several key improvements. Because they present lower physical resistance, advanced rods can now be installed in terrain where only large drill rigs could install before and can quickly be installed in less demanding environments. The modular design of these rods can reduce safety-related accidents during installation. Larger surface areas can improve electrical conductance between the soil and the electrode.

Of particular interest is that advanced driven rods can easily be installed to depths of 20 ft or more depending upon soil conditions. Advanced driven rods are typically driven into the ground with a

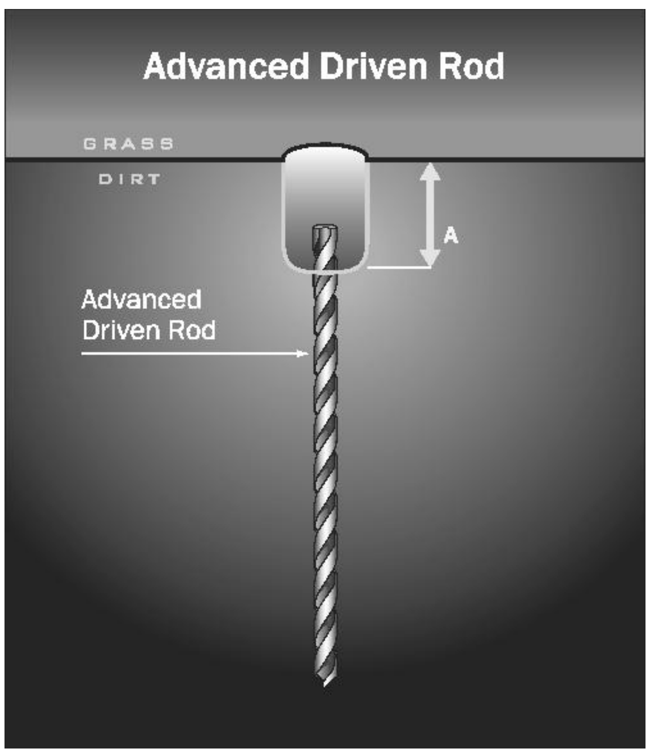


FIGURE 24-3 Advanced driven grounding rod.

standard drill hammer. The tip of an advanced driven rod is typically made of carbide and works in a similar manner to a masonry drill bit, allowing the rod to bore through rock with relative ease. Advanced driven rods are modular in nature and are designed in 5 ft lengths. They have permanent and irreversible connections that enable them to be installed safely while standing on the ground. Typically, a shovel is used to dig down into the ground 18 in before the advanced driven rod is installed. The advanced driven rod falls into the same category as a driven rod and is applicable to the same codes and regulations.

In the extreme northern and southern climates of the planet, frost-heave is a major concern. As frost sets in every winter, unsecured objects buried in the earth tend to be pushed up and out of the ground. Driven grounding rods are particularly susceptible to this action. Anchor plates are often welded to the bottom of the rods to prevent them from being pushed up and out of the earth by frost-heave. This, however, requires that a hole be augured into the earth in order to get the anchor plate into the ground, which can dramatically increase installation costs. Advanced driven rods do not suffer from frost-heave issues and can be installed easily in extreme climes.

24.3.3 Grounding Plates

Grounding plates are typically thin copper plates buried in direct contact with the earth (Fig. 24-4). The *NEC* requires that ground plates have at least 2 ft^2 of surface area exposed to the surrounding soil. Ferrous materials must be at least 0.20 in thick, while nonferrous materials (copper) need only be 0.060 in thick. Grounding plates are typically placed under poles or supplementing counterpoises.

As shown in “A” on Fig. 24-4, grounding plates should be buried at least 30 in below grade level. While the surface area of grounding plates is greatly increased over that of a driven rod, the zone of influence is relatively small as shown in “B.” The *zone of influence* of a grounding plate can be as small as 17 in. This ultra-small zone of influence typically causes grounding plates to have a higher resistance reading than other electrodes of similar mass. Similar environmental conditions that lead to the failure of the driven rod also plague the grounding plate such as, corrosion, aging, temperature, and moisture.

24.3.4 Ufer Ground or Concrete Encased Electrodes

Originally, Ufer grounds were copper electrodes encased in the concrete surrounding ammunition bunkers. In today’s terminology, Ufer grounds consist of any concrete-encased electrode, such as the rebar in building foundations, wire or wire mesh when used for grounding.

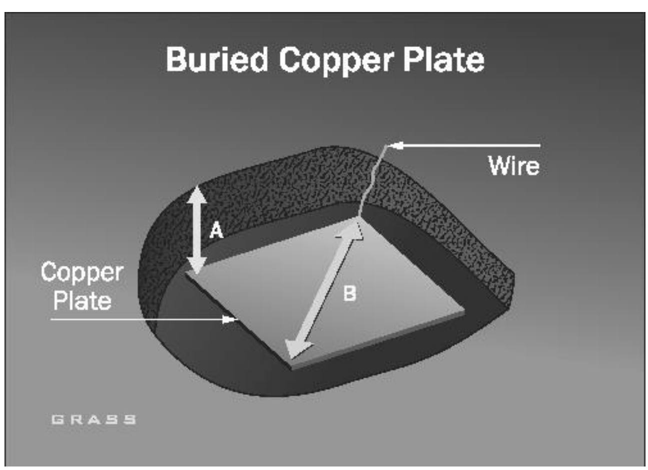


FIGURE 24-4 Buried copper plate.

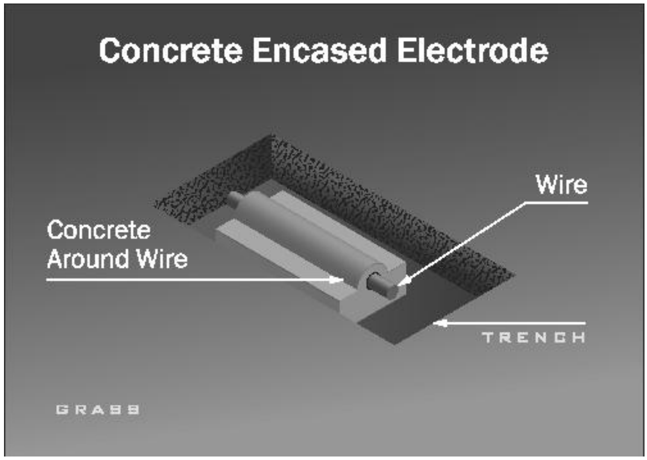


FIGURE 24-5 Concrete encased electrode.

Concrete Encased Electrode. The *NEC* requires that concrete-encased electrodes (Fig. 24-5) use a minimum no. 4 AWG copper wire at least 20 ft in length and encased in at least 2 in of concrete. The advantages of concrete encased electrodes are that they dramatically increase the surface area and amount of contact with the surrounding soil. However, the zone of influence is not increased; therefore, the resistance to ground is typically only slightly lower than the wire would be without the concrete.

Concrete encased electrodes also have some significant disadvantages. When an electrical fault occurs, the electric current must flow through the concrete into the earth. Concrete, by nature, retains a lot of water, which rises in temperature as the electricity flows through the concrete. If the extent of the electrode is not sufficiently great for the total current flowing, the boiling point of the water may be reached, resulting in an explosive conversion of water into steam. Many concrete-encased electrodes have been destroyed after being subjected to relatively small electrical faults. Once the concrete cracks apart and falls away from the conductor, the concrete pieces act as a shield preventing the copper wire from contacting the surrounding soil, resulting in the dramatic increase in the RTG of the electrode.

There are many new products available on the market designed to improve the concrete encased electrodes. The most common are modified concrete products that incorporate conductive materials into the cement mix and are usually carbon. The advantage of these products is that they are fairly effective in reducing the resistivity of the concrete, thus lowering the RTG of the electrode encased. The most significant improvement of these new products is in reducing heat buildup in the concrete during fault conditions, which can lower the chances that steam will destroy the concrete encased electrode.

However, some disadvantages are still evident. Again, these products do not increase the zone of influence and as such the RTG of the concrete encased electrode is only slightly better than what a bare copper wire or driven rod would be in the ground. Also a primary concern regarding enhanced grounding concretes is the use of carbon in the mix. Carbon and copper are of different nobilities and will sacrificially corrode each other over time. Many of these products claim to have buffer materials designed to reduce the accelerated corrosion of the copper caused by the addition of carbon into the mix. However, few independent long-term studies are being conducted to test these claims.

Ufer Ground or Building Foundations. Ufer grounds (Fig. 24-6) or building foundations may be used provided that the concrete is in direct contact with the earth (no plastic moisture barriers); that rebar is at least 0.500 in in diameter; and that there is a direct metallic connection from the service ground to the rebar buried inside the concrete.

This concept is based on the conductivity of the concrete and the large surface area. This will usually provide a grounding system that can handle very high current loads. The primary drawback

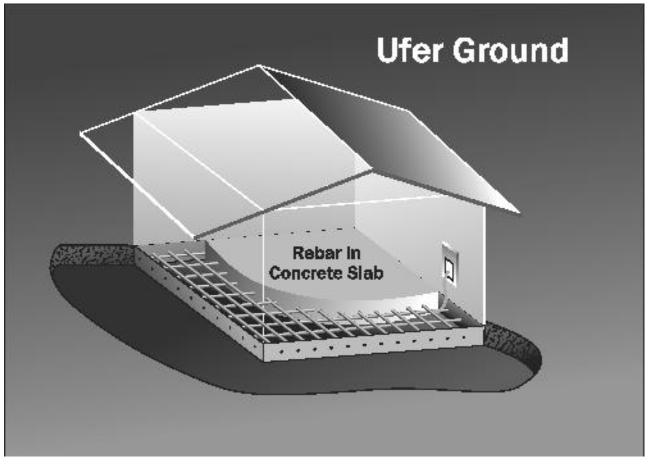


FIGURE 24-6 Building foundation or Ufer.

occurs during fault conditions. If the fault current is too great compared with the area of the rebar system when moisture in the concrete superheats and rapidly expands, cracking the surrounding concrete and threatening the integrity of the building foundation. Another drawback to the Ufer ground is that they are not testable under normal circumstances, as isolating the concrete slab in order to properly perform resistance-to-ground testing is nearly impossible.

The metal frame of a building may also be used as a grounding point, provided that the building foundation meets the above requirements and is commonly done in high-rise buildings. It should be noted that many owners of these high-rise buildings are banning this practice and insisting that tenants run ground wires all the way back to the secondary service locations on each floor. The owners will have already run ground wires from the secondary services back to the primary service locations and installed dedicated grounding systems at these service locations. The goal is to avoid the flow of stray currents that can interfere with the operation of sensitive electronic equipment.

24.3.5 Water Pipes

Water pipes have been used extensively in the past as a grounding electrode. Water pipe connections are not testable and are unreliable due to the use of tar coatings and plastic fittings. City water departments have begun to specifically install plastic insulators in the pipelines to prevent the flow of current and reduce the corrosive effects of electrolysis. Since water pipes are continuous city wide, fault conditions in adjacent neighborhoods could backfeed current into sensitive equipment causing unintentional damage. The *NEC* requires that at least one additional electrode be installed when using water pipes as an electrode. There are several additional requirements including:

- 10 ft of the water pipe is in direct contact with the earth
- Joints must be electrically continuous
- Water meters may not be relied upon for the grounding path
- Bonding jumpers must be used around any insulating joints, pipe, or meters
- Primary connection to the water pipe must be on the street side of the water meter
- Primary connection to the water pipe shall be within 5 ft of the point of entrance to the building

The *NEC* requires that water pipes be bonded to ground, even if water pipes are not used as part of the grounding system.

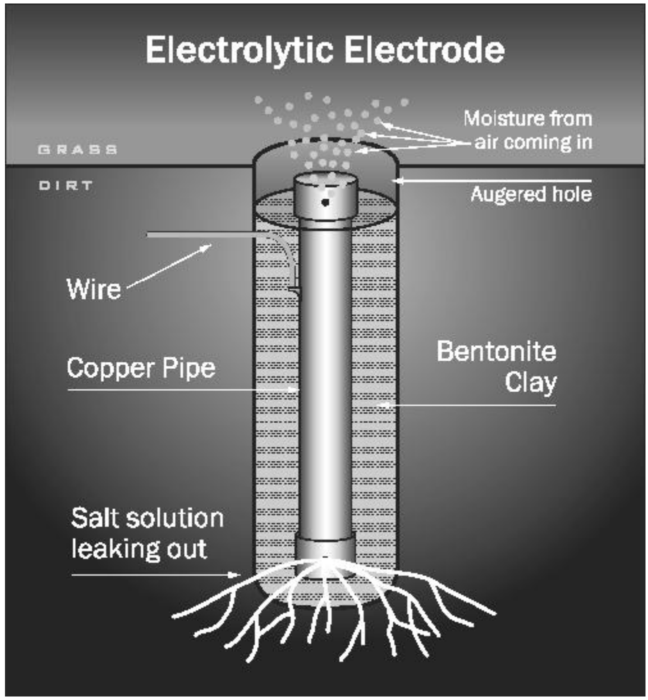


FIGURE 24-7 Electrolytic electrode.

24.3.6 Electrolytic Electrode

The electrolytic electrode (Fig. 24-7) was specifically engineered to eliminate the drawbacks found in other grounding electrodes. This active grounding electrode consists of a hollow copper shaft filled with natural earth salts and desiccants that have a hygroscopic nature to draw moisture from the air. The moisture mixes with the salts to form an electrolytic solution that continuously seeps into the surrounding backfill material, keeping it moist and high in ionic content.

The electrolytic electrode is installed into an augured hole and backfilled with a special highly conductive product. This specialty product should protect the electrode from corrosion and improve its conductivity. The electrolytic solution and the special backfill material work together to provide a solid connection between the electrode and the surrounding soil that is free from the effects of temperature, environment, and corrosion. This active electrode is the only grounding electrode that improves with age. All other electrode types will have a rapidly increasing RTG as the season's change and the years pass. The drawbacks to these electrodes are the cost of installation and the cost of the electrode itself.

24.4 SYSTEM DESIGN AND PLANNING

A grounding design starts with a site analysis, collection of geological data, and soil resistivity of the area. Typically the site engineer or equipment manufacturers specify a RTG number. The NEC states that the resistance-to-ground shall not exceed 25 Ω for a single electrode. However, high technology manufacturers will often specify 3 or 5 Ω depending upon the requirements of their equipment. For sensitive equipment and under extreme circumstances a one ohm specification may sometimes be required.

TABLE 24-2 Surface Materials vs. Resistivity

	Resistivity of sample in Ωm	
Type of surface material	Dry	Wet
Crusher granite w/fines	140×10^6	1,300
Crusher granite w/fines 1.5"	4,000	1,200
Washed granite—pea gravel	40×10^6	5,000
Washed granite 0.75"	2×10^6	10,000
Washed granite 1–2"	1.5×10^6 to 4.5×10^6	5,000
Washed granite 2–4"	2.6×10^6 to 3×10^6	10,000
Washed limestone	7×10^6	2,000 to 3,000
Asphalt	2×10^6 to 30×10^6	10,000 to 6×10^6
Concrete	1×10^6 to 1×10^9	21 to 100

When designing a ground system the difficulty and costs increase exponentially as the target RTG approaches the unobtainable goal of 0 Ω.

24.4.1 Data Collection

Once a need is established, data collection begins. Soil resistivity testing, geological surveys, and test borings provide the basis for all grounding design. Proper soil resistivity testing using the Wenner 4-point method is recommended because of its accuracy. This method will be discussed later in this chapter. Additional data is always helpful and can be collected from the existing ground systems located at the site. For example, driven rods at the location can be tested using the 3-point fall-of-potential method or an induced frequency test using a clamp-on ground resistance meter.

24.4.2 Data Analysis

With all the available data, sophisticated computer programs can provide a soil model showing the resistivity in Ωms and at various layer depths. Knowing at what depth the most conductive soil is located for the site allows the design engineer to model a system to meet the needs of the application.

24.4.3 Grounding Design

Soil resistivity is the key factor that determines the resistance or performance of a grounding system. It is the starting point of any grounding design. As you can see in Figs. 24-10 and 24-11, soil resistivity varies dramatically throughout the world and is heavily influenced by electrolyte content, moisture, minerals, compactness, and temperature (Table 24-2).

24.5 SOIL RESISTANCE TESTING

Soil resistance testing or soil resistivity testing is the process of measuring a volume of soil to determine the conductivity of the soil. The resulting soil resistivity is expressed in Ωm or ohm-centimeter.

Soil Resistivity testing is the single most critical factor in electrical grounding design. This is true when discussing simple electrical design, to dedicated low-resistance grounding systems, or to the far more complex issues involved in GPR studies. Good soil models are the basis of all grounding designs and they are developed from accurate soil resistance testing.

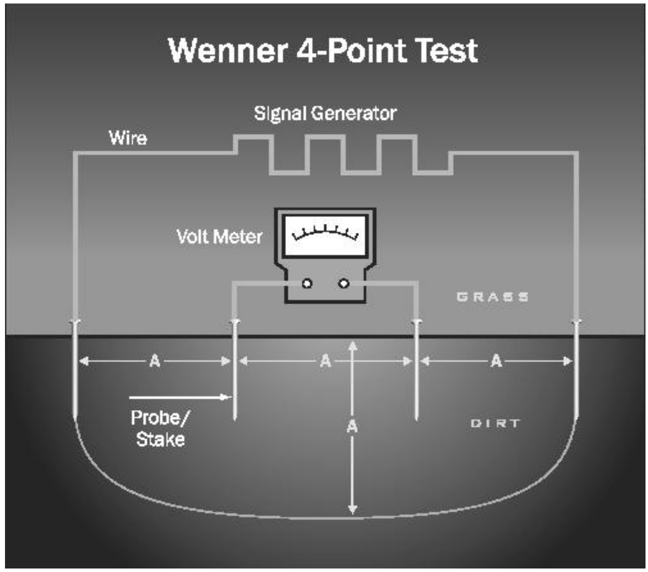


FIGURE 24-8 Wenner 4-point testing pattern.

24.5.1 Wenner Soil Resistivity Test or 4-point Test

The Wenner 4-point method (Fig. 24-8) is by far the most used test method to measure the resistivity of soil. Other methods do exist, such as the General & Schlumberger method, however, they are infrequently used and vary only slightly in how the probes are spaced when compared to the Wenner method.

Electrical resistivity is the measurement of resistance in a unit quantity of a given material. It is expressed in Ωms and represents the resistance measured between two plates covering opposite sides of a 1 m cube. This test is commonly performed at raw land sites, during the design and planning of grounding systems specific to the tested site. The test spaces four probes at equal distances to approximate the depth of the soil to be tested.

Typical spacing will be 1, 1.5, 2, 3, 4.5, 7, 10 ft, etc., with each spacing increasing from the preceding one by a factor of approximately 1.5, up to maximum spacing that is commensurate with the 1 to 3 times the maximum diagonal dimension of the grounding system being designed, resulting in a maximum distance between the outer current electrodes of 3 to 9 times the maximum diagonal dimension of the future grounding system. This is one "traverse" or set of measurements, and is typically repeated, albeit with shorter maximum spacing, several times around the location at right angles and diagonally to each other to ensure accurate readings.

The basic premise of the test is that probes spaced at a 5-ft distance across the earth, will read 5 ft in depth. The same is true if you space the probes 40 ft across the earth, you a weighted average soil resistance down to 40 ft in depth and all points in between. This raw data is usually processed with computer software to determine the actual resistivity of the soil as a function of depth.

Conducting a Wenner 4-point (or four-pin) Test. Figure 24-9 shows how to take one "traverse" or set of measurements. As the 4-point indicates, the test consists of four pins that must be inserted into the earth. The outer two pins are called the current probes, C1 and C2. These are the probes that inject current into the earth. The inner two probes are the potential probes, P1 and P2. These are the probes that take the actual soil resistance measurement.

In this test, a probe C1 is driven into the earth at the corner of the area to be measured. Probes P1, P2, and C2 are driven at 5, 10, and 15 ft respectively from rod C1 in a straight line to measure

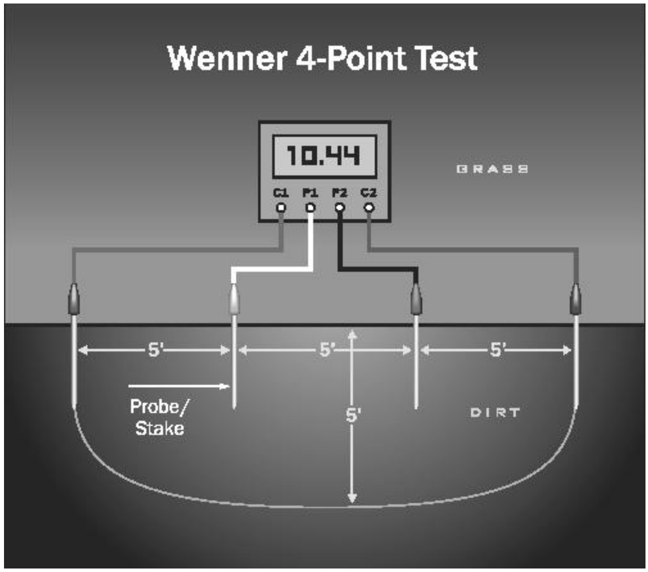


FIGURE 24-9 Wenner 4-point test setup.

the soil resistivity from 0 to 5 ft in depth. C1 and C2 are the outer probes and P1 and P2 are the inner probes. At this point, a known current is applied across probes C1 and C2, while the resulting voltage is measured across P1 and P2. Ohm's law can then be applied to calculate the measured resistance.

Probes C2, P1, and P2 can then be moved out to 10, 20, and 30 ft spacing to measure the resistance of the earth from 0 to 10 ft in depth. Continue moving the three probes (C2, P1, and P2) away from C1 at equal intervals to approximate the depth of the soil to be measured. Note that the performance of the electrode can be influenced by soil resistivities at depths that are considerably deeper than the depth of the electrode, particularly for extensive horizontal electrodes, such as water pipes, building foundations or grounding grids.

Soil Resistance Meters. There are basically two types of soil resistance meters: low- and high-frequency models. Both meter types can be used for 4-point and 3-point testing, and can even be used as standard (2-point) voltmeter for measuring common resistances.

Care should always be given when selecting a meter, as the electronics involved in signal filtering are highly specialized. Electrically speaking, the earth can be a noisy place. Overhead power lines, electric substations, railroad tracks, various signal transmitters, and many other sources contribute to signal noise found in any given location. Harmonics, 60 Hz background noise, and magnetic field coupling can confound the measurement signal resulting in apparent soil resistivity readings that are larger by an order of magnitude, particularly with large spacings. Selecting equipment with electronic packages capable of discriminating between these signals is critical.

High-Frequency meters typically operate at 128 pulses per second or other pulse rates except 60. These high-frequency meters typically suffer from the inability to generate sufficient voltage to handle long traverses and generally should not be used for probe spacings greater than 100 ft. Furthermore, the high-frequency signal flowing in the current lead induces a noise voltage in the potential leads, which cannot be completely filtered. This noise becomes greater than the measured signal as the soil resistivity decreases and the pin spacing increases. High-frequency meters are less expensive than their low-frequency counterparts and are by far the most common meter used in soil resistivity testing.

TABLE 24-3 Soil Types vs. Resistivity Chart

Soil types or type of earth	Average resistivity in Ωm
Bentonite	2 to 10
Clay	2 to 100
Wet organic soils	10 to 100
Moist organic soils	100 to 1,000
Dry organic soils	1,000
Sand and gravel	50 to 1,000
Surface limestone	100 to 10,000
Limestone	5 to 4,000
Shale's	5 to 100
Sandstone	20 to 2,000
Granites, basalts, etc.	1,000
Decomposed gneisses	50 to 500
Slates, etc.	10 to 100

Low-Frequency meters, which actually generate low frequency pulses (on the order of 0.5 to 2.0 seconds per pulse), are the preferred equipment for soil resistivity testing, as they do away with the induction problem from which the high-frequency meters suffer. However they can be very expensive to purchase. Depending upon the equipment's maximum voltage, low-frequency meters can take readings with extremely large probe spacings and often many thousands of feet in distance. Typically, the electronics filtering packages offered in low-frequency meters are superior to those found in high-frequency meters. Caution should be taken to select a reputable manufacturer.

Data Analysis. Once all the resistance data are collected, formula 24-3 can be applied to calculate the apparent soil resistivity in Ωms. For example, if an apparent resistance of 4.5 Ω is at 40-ft spacing, the soil resistivity in Ωms would be 344.7. One refers to "apparent" resistivity, because this does not correspond to the actual resistivity of the soil. This raw data must be interpreted by suitable methods in order to determine the actual resistivity of the soil (Table 24-3).

Soil Resistivity
4-Point Data Interpretation

$$\rho = 1.915AR$$

$$\rho = 1.915(40)(4.5)$$

$$\rho = \frac{4\pi AR}{1 + \frac{2A}{\sqrt{(A^2 + 4B^2)} - \sqrt{(A^2 + B^2)}}}$$

ρ = Resistivity　　A = Spacing of Probes

B = Depth of Probes　　R = Resistance (reading from meter)

If $A > 20B$, then $\rho = 2\pi AR = 1.915\ AR$

Shallow Depth Readings. Shallow depth readings, as little as 6 in deep are exceedingly important for most, if not all, grounding designs. The soil resistivity readings are actually weighted averages of the soil resistivity from the earth's surface down to the specified distance and include all the shallow

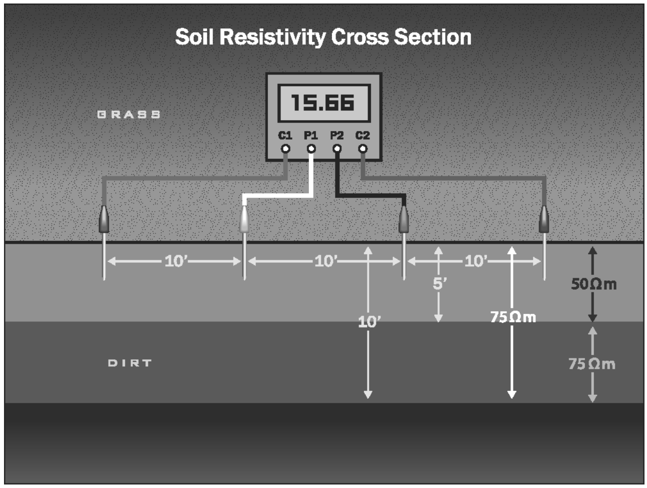


FIGURE 24-10 Importance of shallow readings.

resistance readings above it. The trick in developing the final soil model is to pull out the actual resistance of the soil at depth, and that requires subtracting the top layers from the deep readings.

Figure 24-10 demonstrates how the shallowest readings impact deeper ones below it. If the 5-ft reading shows 50-Ωm soil, and the 10-ft reading shows 75-Ωm soil, the actual soil resistance from 5 to 10 ft is 100-Ωm. The point here is to illustrate a concept that precomputed curves or computer software is needed to properly interpret the data. The same follows true for larger pin spacings. The shallowest readings are used over and over again in determining the actual resistance at depth.

Shallow-depth readings of 6 in, 1 ft, 1.5 ft, 2 ft, and 2.5 ft are important for grounding design. Because grounding conductors are typically buried at 1.5 to 2.5 ft below the surface of the earth. To accurately calculate how those conductors will perform at these depths, shallow soil readings must be taken. These shallow readings become even more important when engineers calculate GPR and step and touch voltages.

It is critical that the measurement probes and current probes be inserted into the earth to the proper depth for shallow soil resistivity readings. If the probes are driven too deep, then it can be difficult to resolve the resistivity of the shallow soil. A rule of thumb is that the penetration depth of the potential probes should be no more than 10% of the pin spacing, whereas the current probes must not be driven more than 30% of the pin spacing.

Deep Readings. Often, the type of meter used determines the maximum depth or spacing that can be read. A general guideline is that high-frequency soil resistivity meters are good for no more than 100-ft pin spacings, particularly in low resistivity soils. For greater pin spacings, low-frequency soil resistivity meters are required. They can generate the required voltage needed to push the signal through the soil at deep distances and detect a weak signal, free of induced voltage from the current injection leads.

24.5.2 Test Location

Soil resistivity testing should be conducted as close to the proposed grounding system as possible, taking into consideration the physical items that may cause erroneous readings. There are two issues that may cause poor quality readings:

1. Electrical interference causing unwanted signal noise to enter the meter.
2. Metallic objects “short-cutting” the electrical path from probe to probe. The rule of thumb here is that a clearance equal to the pin spacing should be maintained between the measurement traverse and any parallel buried metallic structures.

Testing in the vicinity of the site in question is obviously important; however, it is not always practical. Many electric utility companies have rules regarding how close the soil resistivity test must be in order to be valid. The geology of the area also plays into the equation as dramatically different soil conditions may exist only a short distance away.

When left with little room or poor conditions in which to conduct a proper soil resistivity test, one should use the closest available open field with as similar geological soil conditions as possible.

24.6 TESTING OF EXISTING GROUNDING SYSTEMS

The measurement of ground resistance for an earth electrode system is very important. It should be done when the electrode is first installed, and then at periodic intervals thereafter. This ensures that the RTG does not increase over time. There are two methods for testing an existing earth-electrode system. The first is the 3-point or fall-of-potential method and the second is the Induced Frequency test or clamp on method. The 3-point test requires complete isolation from the power utility. Not just power isolation, but also removal of any neutral or other such ground connections extending outside the grounding system. This test is the most suitable test for large grounding systems and is also suitable for small electrodes. The induced frequency test can be performed while power is on and actually requires the utility to be connected to the grounding system under test. This test is accurate only for small electrodes, as it uses frequencies in the kilohertz range, which see long conductors as inductive chokes and therefore do not reflect the 60 Hz resistance of the entire grounding system.

24.6.1 Fall-of-Potential Method or 3-point Test

The 3-point or fall-of-potential method (Fig. 24-11) is used to measure the RTG of existing grounding systems. The two primary requirements to successfully complete this test are the ability to isolate the grounding system from the utility neutral and knowledge of the diagonal length of the grounding system (i.e., a 10-ft × 10-ft grounding ring would have a 14 ft diagonal length). In this test, a short probe, referred to as probe Z, is driven into the earth at a distance of 10 times (10×) the

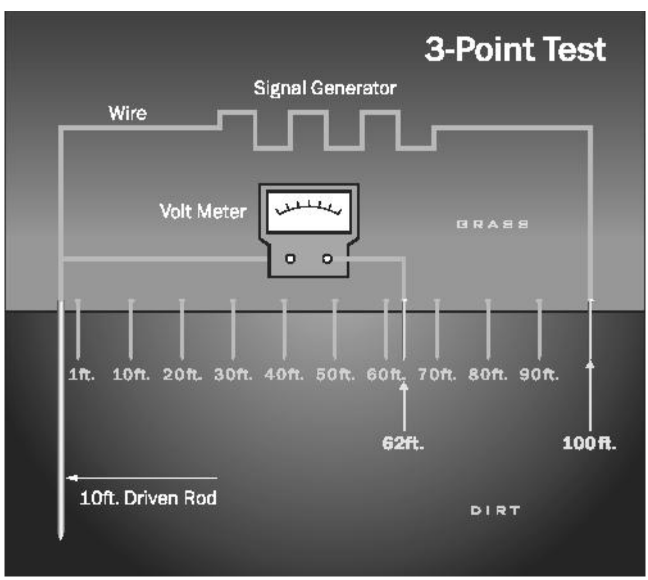


FIGURE 24-11 3-point test method.

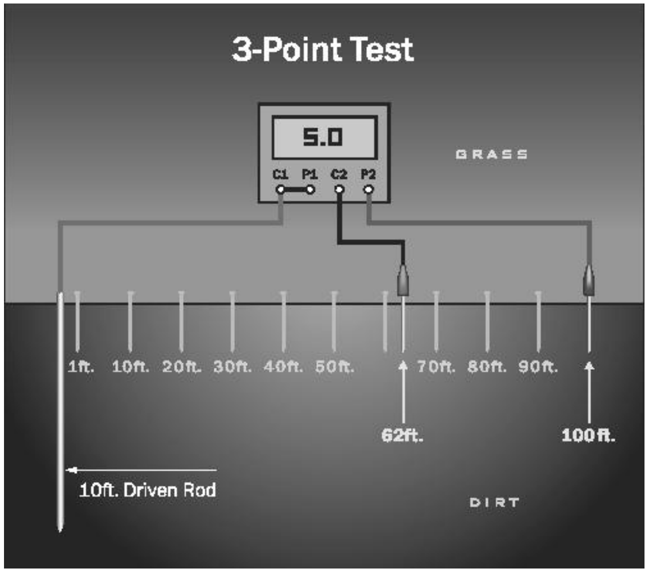


FIGURE 24-12 3-point test setup.

diagonal length of the grounding system (rod X). A second probe (Y) is placed in-line at a distance from rod X equal to the diagonal length of the grounding system.

At this point, a known current is applied across X and Z, while the resulting voltage is measured across X and Y. Ohm's Law can then be applied ($R = V/I$) to calculate the measured resistance. Probe Y is then moved out to a distance of 2x the diagonal length of the grounding system, in-line with X and Z, to repeat the resistance measurement at the new interval. This will continue, moving probe Y out to 3x, 4x, … 9x the diagonal length to complete the 3-point test (Fig. 24-12) with a total of nine resistance measurements.

Graphing & Evaluation. The 3-point test is evaluated by plotting the results as data points with the distance from rod X along the *x*-axis and the resistance measurements along the *y*-axis to develop a curve. Roughly midway between the center of the electrode under test and the probe Z, a plateau or "flat spot" should be found, as shown in the graph. The resistance of this plateau (actually, the resistance measured at the location 62% from the center of the electrode under test, if the soil is perfectly homogeneous) is the RTG of the tested grounding system.

Invalid Tests. If no semblance of a plateau is found and the graph is observed to rise steadily, the test is considered invalid (Fig. 24-13). This can be due to the fact that probe Z was not placed far enough away from rod X, and can usually indicate that the diagonal length of the grounding system was not determined correctly. If the graph is observed to have a low plateau that extends the entire length and only rises at the last test point, then this also may be also considered invalid. This is because the utility or telecom neutral connection remains on the grounding system.

24.6.2 Induced Frequency Testing or Clamp-On Testing

The induced frequency testing (Fig. 24-14) or commonly called the "clamp-on" test is one of the newest test methods for measuring the RTG of a grounding system or electrode. This test uses a special transformer to induce an oscillating voltage (often 1.7 kHz) into the grounding system. Unlike the 3-point test which requires the grounding system to be completely disconnected and isolated

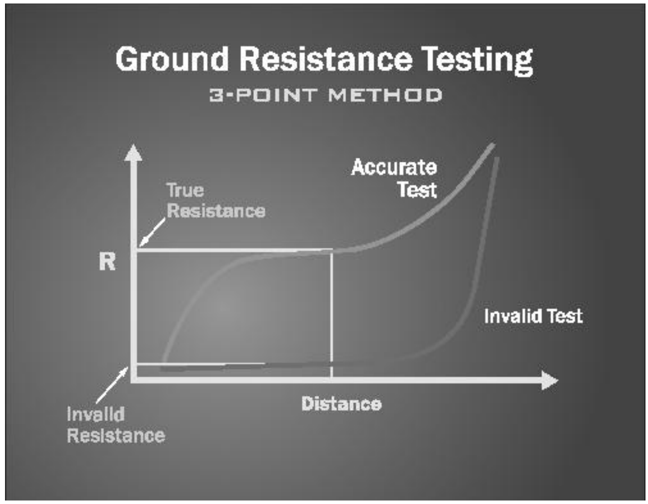


FIGURE 24-13 Comparison of valid vs. invalid 3-point test.

before testing, this method requires that the grounding system under test be connected to the electric utilities (or other large grounding system such as from the telephone company) grounding system (typically via the neutral return wire) to provide the return path for the signal. This test is the only test that can be used on "live" or "hot" systems. However, there are some limitations, including:

1. The amount of current running through the tested system must be below the equipment manufacturers limits.
2. The test signal must be injected at the proper location, so that the signal is forced through the grounding system and into the earth.
3. This instrument actually measures the sum of the resistance of the grounding system under test and the impedance of the utility neutral grounding, including the neutral wiring. Due to the high frequency used, the impedance of the neutral wiring is nonnegligible and can be greater than the

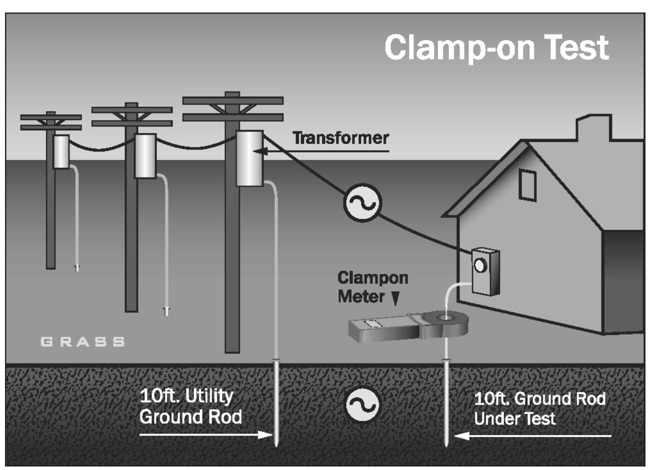


FIGURE 24-14 Induced frequency test diagram.

ground resistance of a very low resistance grounding system, which can therefore not be measured accurately.

4. The ground resistance of a large grounding system at 60 Hz can be significantly lower than at 1.7 kHz.

Many erroneous tests have been conducted where the technician only measured metallic loops and not the true RTG of the grounding system. The veracity of the induced frequency test has been questioned due to testing errors; however, when properly applied to a small to medium sized, self-standing grounding system, this test is rapid and reasonably accurate.

Test Application. The proper use of this test method requires the utility neutral to be connected to a wye-connected transformer. The oscillating voltage is induced into the grounding system at a point where it will be forced into the soil and return through the utility neutral. Extreme caution must be taken at this point as erroneous readings and mistakes are often made. The most common of these occur when clamping on or inducing the oscillating voltage into the grounding system at a point where a continuous metallic path exists back to the point of the test. This can result in a continuity test being performed rather than a ground resistance test. Understanding the proper field application of this test is vital to obtaining accurate results. The induced frequency test can test grounding systems that are in use and does not require the interruption of service to take measurements.

Ground Resistance Monitoring. Ground resistance monitoring is the process of automated timed and/or continuous RTG measurement. These dedicated systems use the induced frequency test method to continuously monitor the performance of critical grounding systems. Some models may also provide automated data reporting. These new meters can measure RTG and the current that flows on the grounding systems that are in use. Another benefit is that it does not require interruption of the electrical service to take these measurements.

24.7 GROUND POTENTIAL RISE

Ground potential rise is a phenomenon that occurs when large amounts of electricity enter the earth. This is typically caused when substations or high-voltage towers fault, or when lightning strikes occur (fault current). When currents of large magnitude enter the earth from a grounding system, not only will the grounding system rise in electrical potential, but so will the surrounding soil as well (Fig. 24-15).

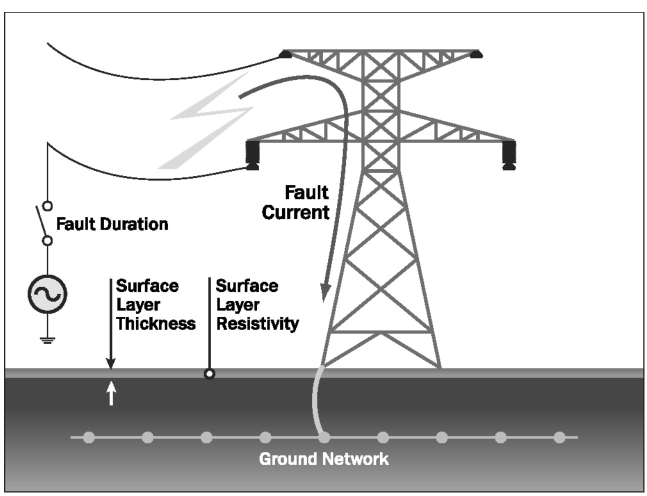


FIGURE 24-15 Electric fault at a transmission tower.

The voltages produced by a GPR event can be hazardous to both personnel and equipment. As described earlier, soil has resistance which will allow an electrical potential gradient or voltage drop to occur along the path of the fault current in the soil. The resulting potential differences will cause currents to flow into any and all nearby grounded conductive bodies, including concrete, pipes, copper wires, and people.

Ground Potential Rise Definitions. GPR as defined in IEEE Std. 367 is the product of a ground electrode impedance, referenced to remote earth, and the current that flows through that electrode impedance.

GPR as defined by IEEE Std. 80-2000 as the maximum electrical potential a substation grounding grid may attain, relative to a distant grounding point assumed to be at the remote earth potential. This voltage, GPR, is equal to the maximum grid current times the grid resistance.

Ground Potential Rise events are a concern wherever electrical currents of large magnitude flow into the earth. This can be at a substation, high-voltage tower or pole, or a large transformer. In cases where a GPR event may be of special concern, grounding precautions are required to ensure personnel and equipment safety.

Electrical potentials in the earth drop abruptly around the perimeter of a grounding system, but do not drop to zero. In fact, in a perfectly homogeneous soil, soil potentials are inversely proportional to the distance from the center off the grounding system, once one has reached a distance that is a small number of grounding system dimensions away. The formula is as follows:

$$\text{Earth potential} = \text{soil resistivity} \times \text{current}/(2 \times PI \times \text{distance})$$

Where earth potential is in volts, soil resistivity is in Ωms, current is the current flowing into the soil from the grounding system, in amperes, *PI* is 3.14159 … and distance is in meters. Probably the most commonly noted GPR event involves the death of cows in a field during a lightning strike. Imagine lightning striking the center of an open field where cows are standing. The current injected into the earth flows radially away from the strike point, in all directions, creating voltage gradients on the surface of the earth, also in a radial direction. All the cows facing the lightning strike would have their fore hooves closer to the strike point than their rear hooves. This would result in a difference of potential between their fore and rear legs, causing current to flow through their bodies, including the heart area, and killing the cow. On the other hand, those cows with their flanks turned towards the lighting strike would have a greater chance of surviving, as the distance between their fore legs and therefore the voltage applied between them, would be relatively small, resulting in a lesser current flow.

GPR studies are typically conducted on substations and high-voltage towers. Substations have relatively large grounding areas, especially when compared to high-voltage towers and poles. Towers and poles represent by far the most potentially dangerous and difficult GPR situations to handle and are often not protected, unless they are located in high exposure areas or have equipment installed at ground level at which service personnel might be required to work.

24.7.1 Ground Potential Rise Analysis

The primary purpose of a GPR study is to determine the level of hazard associated with a given high-voltage location for personnel and/or equipment. When the degree of hazard is identified the appropriate precautions must be made to make the site safe. To do this, the engineer must identify what the minimum grounding system for each location will be. The engineer must also take into consideration all local and federal guidelines including utility companies and other requirements.

For example, many utility companies require at a minimum that a simple ground ring be installed at least 18 in below ground and 3-ft from the perimeter of all metal objects. This ground ring is also referred to as a counterpoise.

Once the minimum grounding system is identified, the engineer can run a GPR analysis and identify the extent of any electrical hazards.

Typically items reported in a GPR study will include the following: The square footage and size of the proposed grounding grid, RTG of the proposed grounding system, the estimated fault current that would be seen at the site, voltage rise of the GPR (in volts) at the site, 300-Vpeak line, the *X*/*R* ratio, and the fault clearing time in seconds. Touch and step voltages are usually computed as well, as these are the primary indicators of safety.

The grounding engineer needs three pieces of information to properly conduct a GPR Study:

1. Soil resistivity data
2. Site drawings with the proposed construction
3. Electrical data from the utility company

Soil Resistivity Data. The soil resistivity data should include apparent resistivity readings at pin spacings ranging from 0.5 or 1 ft to as many as three grounding grid diagonals, if practical. Touch and step voltages represent the primary concern for personnel safety. Understanding the characteristics of the soil at depths ranging from immediately underfoot to one or more grid dimensions is required for a cost-effective and safe grounding system to be designed.

Site Drawings. The proposed site drawings should show the layout of the high-voltage tower or substation, and any additional construction for new equipment that may be occurring on the site, including fencing and gate radius. Incoming power and telco runs should also be included. In the case of high-voltage towers, the height and spacing of the conductors carried on the tower, and any overhead ground wires that may be installed on the tower, need to be detailed during the survey. This information is needed to properly address all the touch and step voltage concerns that may occur on the site.

Electric Utility Data. The electric utility company needs to provide electrical data regarding the tower or substation under consideration. This data should include the name of the substation or the number of the tower, the voltage level, the subtransient *X*/*R* ratio, and the clearing times. In the case of towers, the line names of the substations involved, the amount of current contributed by each substation in the event of a fault, and the type and positions of the overhead ground wires, if any, with respect to the phase conductors installed on each tower or pole. If overhead ground wires are present, tower or pole ground resistances along the line are of interest as well, be they measured, average or design values.

This information is important, as high-voltage towers have small ground area, yet handle very large amounts of electricity. Knowing if a tower has an overhead ground wire is important, because the overhead wire will carry away a percentage of the current, which will depend on the overhead ground wire type and ground resistances of adjacent towers, to other towers in the run, reducing the GPR event. Additionally, towers with overhead ground wires tend to have shorter clearing times. The same holds for substations—overhead ground wires on transmission lines and neutral wires on distribution lines can significantly reduce the magnitude of fault current that flows into the substation grounding system during fault conditions.

The following information is required from the utility company:

1. Phase-to-ground fault current contributed by each power line circuit
2. Fault clearing time
3. Line voltage
4. Subtransient *X*/*R* ratio
5. The make/type/number of overhead ground wires on each tower/pole line and position with respect to the phase conductors
6. Ground wire continuity and bonding configuration back to the tower and substation

7. The average distance from tower-to-tower and tower-to-substation

8. Typical tower/pole ground resistance: measured or design values

As-built drawings are often acquired and are useful for towers with existing grounding systems. They are also useful in the case of modifications and upgrades to existing substations, which will have extensive grounding systems already installed.

24.7.2 Personnel Safety during Ground Potential Rise Events

The grounding engineer will be required to develop safety systems to protect any personnel working where GPR hazards are known to exist. Federal law mandates that all known hazards must be eliminated from the work place for the safety of workers. It is the engineer's choice on which voluntary standards to apply in order to comply with the law. Federal law 29 CFR 1910.269 specifically states that step and touch potentials must be eliminated on transmission and distribution lines that include any related communication equipment.

Substations are always considered workplaces and step and touch potentials must be eliminated. Transmission and distribution towers or poles are not always considered work places and therefore are often exempt from these requirements. Take, for example, a lonely tower on a mountain side or in the middle of the desert—these towers are not typically considered workplaces. However, any high-voltage tower or pole becomes a workplace as soon as equipment is installed that is not related to the electric utility company and requires outside vendors to support the new equipment. Cellular telecommunications, environmental monitoring, and microwave relay equipment are good examples of equipment that, when installed on a high-voltage tower, turns the tower into a workplace. This would make the elimination of step and touch potentials required.

Hazardous Voltages. Fibrillation current is the amount of electricity needed to cause cardiac arrest, from which recovery will not spontaneously occur, in a person and is a value based on statistics. IEEE Std. 80-2000 provides a method to determine the pertinent value of fibrillation current for a safety study, along with a good explanation of how it is derived. Many different methods exist for calculating fibrillation current; however the 50kg-IEEE method is the most commonly used in North America. The formula used shows that the fibrillation current level is inversely proportional to the square root of the fault duration; however, it must be increased by a correction factor, based on the subtransient X/R ratio, which can be quite large for shorter fault durations. If personnel working at a site during fault conditions experience voltages that will cause a current less than the fibrillation current to flow in their bodies, then they are considered safe. If a worker will experience a greater voltage than is acceptable, additional safety precautions must be taken.

The subtransient X/R ratio at the site of the fault is important in calculating the acceptable fibrillation current and to determine the maximum allowable step and touch voltages that can occur at any given site.

Fault duration is a very necessary piece of data for properly calculating step and touch potentials. The fault duration is the amount of time required for the power company to shut off the current in the event of a fault.

Ultimately the engineer must determine two things:

1. The site-specific maximum allowable voltage that a person can safely withstand

2. The actual voltages that will be experienced at the site during a fault

Each site will have different levels of voltages for both of the above. Unfortunately, we cannot simply say that a human being can withstand X-level of voltages and use that value all the time, since this voltage is determined by the surface layer resistivity, the fault duration and the subtransient X/R ratio. Additionally as each site has different fault durations and different soil conditions, it is critical that calculations be made for each and every possible fault location.

Step Potential. When a fault occurs at a tower or substation, the current will enter the earth. Based on the distribution of varying resistivity in the soil (typically, a horizontally layered soil is assumed) a corresponding voltage distribution will occur. The voltage drop in the soil surrounding the grounding system can present hazards for personnel standing in the vicinity of the grounding system. Personnel "stepping" in the direction of the voltage gradient could be subjected to hazardous voltages.

In the case of step potentials, electricity will flow if a difference in potential exists between the two legs of a person (Fig. 24-16). Calculations must be performed that determine how great the tolerable step potentials are and then compare those results to the step voltages expected to occur at the site.

Hazardous step potentials can occur a significant distance away from any given site. The more current that is pumped into the ground, the greater the hazard. Soil resistivity and layering plays a major role in how hazardous a fault occurring on a specific site may be. High soil resistivities tend to increase step potentials. A high resistivity top layer and low resistivity bottom layer tends to result in the highest step voltages close to the ground electrode; the low resistivity bottom layer draws more current out of the electrode through the high resistivity layer, resulting in large voltage drops near the electrode. Further from the ground electrode, the worst case scenario occurs when the soil has conductive top layers and resistive bottom layers; in this case, the fault current remains in the conductive top layer for much greater distances away from the electrode.

Fault clearing time is an important factor to consider as well. The more time it takes the electric utility company to clear the fault, the more likely it is for a given level of current to cause the human heart to fibrillate.

An important note to remember is that most power companies use automated reclosers. In the event of a fault, the power is shut off and then automatically turned back on. This is done in case the faults occurred due to an unfortunate bird that made a poor choice in where to rest, or dust that may have been burned off during the original fault. A few engineers believe that fibrillation current for step potentials must be far greater than touch potentials, as current will not pass through any vital organs in the former case. This is not always true as personnel that receive a shock due to step potentials may fall to the ground, only to be hit again, before they can get up, when the automatic reclosers activate.

Touch Potentials. When a fault occurs at a tower or substation the current will pass through any metallic object and enter the earth. Those personnel "touching" an object in the vicinity of the GPR

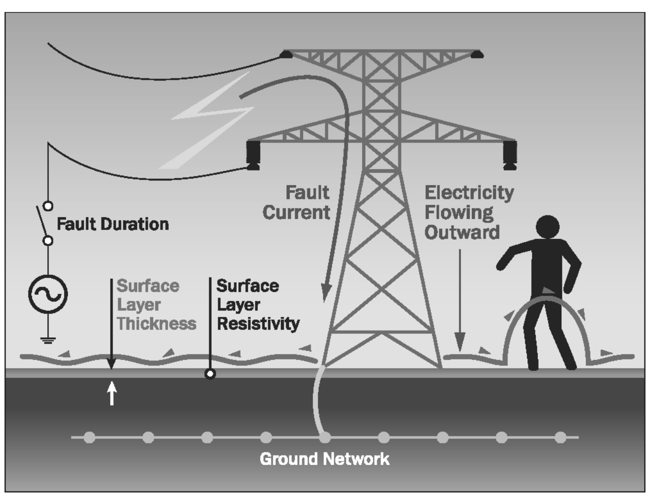


FIGURE 24-16 Step potential at a transmission tower.

will be subjected to these voltages which may be hazardous. For example if a person happens to be touching a high-voltage tower leg when a fault occurs, the current would travel down the tower leg into the persons hand and through vital organs of the body (Fig. 24-17). It would then continue on its path and exit out through the feet and into the earth. Careful analysis is required to determine the acceptable fibrillation currents that can be withstood by the body if a fault were to occur.

Engineering standards use a 1-m (3.28 ft) reach distance for calculating touch potentials. A 2-m (6.54 ft) reach distance is used when two or more objects are inside the GPR event area. For example, a person could be outstretching both arms and touching two objects at once such as a tower leg and a metal cabinet. Occasionally, engineers will use a 3-m distance to be particularly cautious, as they assume someone may be using a power tool with a power cord 3 m in length.

The selection of where to place the reference points used in the touch potential calculations are critical in getting an accurate understanding of the level of hazard at a given site. The actual calculation of touch potentials uses a specified object (such as a tower leg) as the first reference point. This means that the further away from the tower the other reference point is located, the greater the difference in potential. If you can imagine a person with incredibly long arms touching the tower leg and yet standing many dozens of feet away, you would have a huge difference in potential between the feet and the tower. Obviously, this example is not possible—this is why setting where and how far away the reference points used in the touch calculation is so important and why the 1-m rule has been established.

Mitigating Step and Touch Potential Hazards. Mitigating step and touch potential hazards is usually accomplished through one or more of the following three main techniques:

1. Reduction in the RTG of the grounding system
2. Proper placement of ground conductors
3. The addition of resistive surface layers

Understanding the proper application of these techniques is the key to reducing and eliminating any GPR hazards. Only through the use of highly sophisticated three-dimensional electrical simulation software that can model soil structures with multiple layers and finite volumes of different materials, can the engineer accurately model and design a grounding system that will safely handle high-voltage electrical faults.

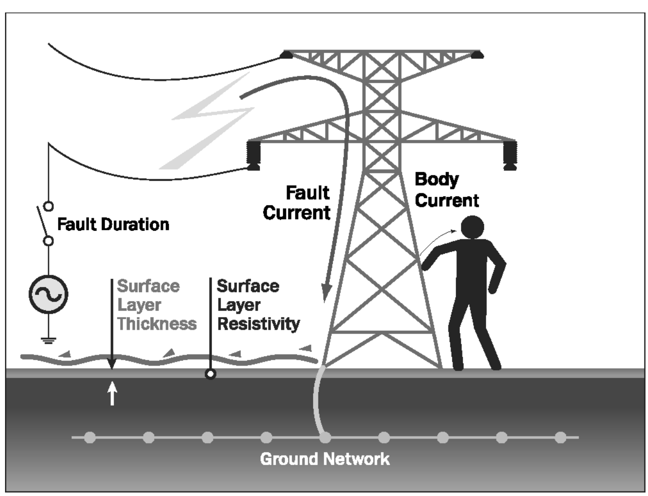


FIGURE 24-17 Touch potential at a transmission tower.

Reduce the Resistance-to-Ground. Reducing the RTG of the site is often the best way to reduce the negative effects of any GPR event, where practical. The GPR is the product of the fault current flowing into the grounding system times the RTG of the grounding system. Thus, reducing the RTG will reduce the GPR to the degree that the fault current flowing into the grounding system does increase in response to the reduced RTG. For example, if the fault current for a high-voltage tower is 5000 A and the RTG of the grounding system is 10-Ω, the GPR will be 50,000 V. If we reduce the RTG of the grounding system down to 5-Ω and the fault current increases to 7000 A as a result, then the GPR will become 35,000 V.

As seen in the example above, the reduction of the RTG can have the effect of allowing more current to flow into the earth at the site of the fault, but will always result in lower GPR values and touch and step voltages at the fault location. On the other hand, further away from the fault location, at adjacent facilities not connected to the faulted structure, the increase in current into the earth will result in greater current flow near these adjacent facilities and therefore an increase in the ground potential rise, touch voltages and step voltages at these facilities. Of course, if these are low to begin with, an increase may not represent a problem, but there are cases in which a concern may exist. Reducing the RTG can be achieved by any number of means as discussed earlier in this chapter.

Proper Placement of Ground Conductors. A typical specification for ground conductors at high-voltage towers or substations is to install a ground loop around all metallic objects, connected to the objects; keep in mind that it may be necessary to vary the depth and/or distance that ground loops are buried from the structure in order to provide the necessary protection. Typically these ground loops require a minimum size of 2/0 AWG bare copper conductor buried in direct contact with the earth and 3-ft from the perimeter of the object, 18 in below grade. The purpose of the loop is to minimize the voltage between the object and the earth surface where a person might be standing while touching the object—that is, to minimize touch potentials.

It is important that all metallic objects in a GPR environment be bonded to the ground system to eliminate any difference in potentials. It is also important that the resistivity of the soil as a function of depth be considered in computed touch and step voltages and in determining at what depth to place conductors. For example, in a soil with a dry, high-resistivity surface layer, conductors in this layer will be ineffective; a low resistivity layer beneath that one would be the best location for grounding conductors. On the other hand, if another high resistivity layer exists further down, long ground rods or deep wells extending into this layer will be ineffectual.

It is sometimes believed that placing horizontal grounding loop conductors very close to the surface results in the greatest reduction in touch potentials. This is not necessarily so, as conductors close to the surface are likely to be in drier soil, with a higher resistivity, thus reducing the effectiveness of these conductors. Furthermore, while touch potentials immediately over the loop may be reduced, touch potentials a short distance away may actually increase due to the decreased zone of influence of these conductors. Finally, step potentials are likely to increase at these locations; indeed, step potentials can be a concern near conductors that are close to the surface, particularly at the perimeter of a grounding system. It is common to see perimeter conductors around small grounding systems buried to a depth of 3-ft below grade, in order to address this problem.

Resistive Surface Layers. One of the simplest methods of reducing step and touch potential hazards is to wear electric hazard shoes. When dry, properly rated electric hazard shoes have millions of ohms of resistance in the soles and are an excellent tool for personnel safety. On the other hand, when these boots are wet and dirty, current may bypass the soles of the boots in the film of material that has accumulated on the sides of the boot. A wet leather boot can have a resistance on the order of 100 Ω. Furthermore, it cannot be assumed that the general public, who may have access to the outside perimeter of some sites, will wear such protective gear.

Another technique used in mitigating step and touch potential hazards is the addition of more resistive surface layers. Often a layer of crushed rock is added to a tower or substation to provide a layer of insulation between personnel and the earth. This layer reduces the amount of current that can flow through a given person and into the earth. Weed control is another important factor, as plants become energized during a fault and can conduct hazardous voltages into a person. Asphalt is an excellent alternative, as it is far more resistive than crushed rock, and weed growth is not a problem. The addition of resistive surface layers always improves personnel safety during a GPR event.

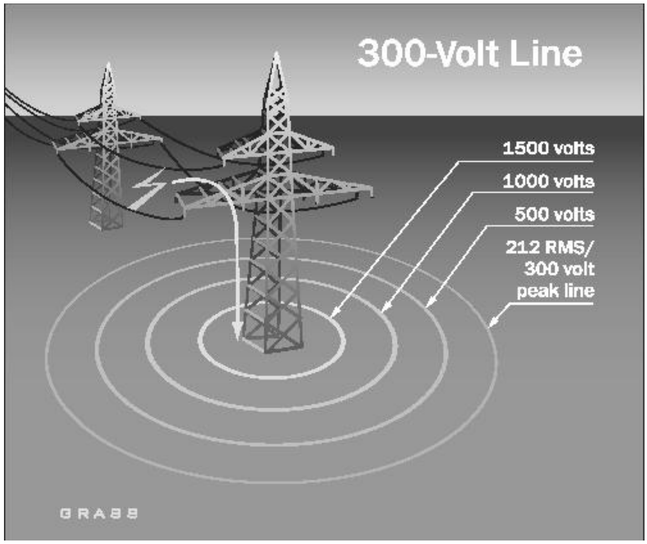


FIGURE 24-18 300-V line at a transmission tower.

Telecommunications in High-Voltage Environments. When telecommunications lines are needed at a high-voltage site, special precautions are required to protect switching stations from unwanted voltages. Installing wires in a substation or on a tower could present a hazardous situation and therefore certain precautions are required.

Industry standards regarding these precautions and protective requirements are covered in IEEE Standard's 387, 487, and 1590. These standards require that a GPR study be conducted so that the 300-V peak line can be properly calculated.

To protect the telephone switching stations, telecommunication standards require that fiber-optic cables be used instead of copper wires. A copper-to-fiber conversion box must be located outside the GPR event area at a distance in excess of the 300-V peak or 212-V RMS line (Fig. 24-18). This is known in the industry as the "300-V line." This means that based on the calculation results, copper wire from the telecommunications company may not come any closer than the 300-V peak distance. This is the distance where copper wire must be converted to fiber-optic cable. This can help prevent any unwanted voltages from entering the phone companies' telecommunications network.

The current formulae for calculating the 300-volt line, as listed in the standards, has led to misinterpretation and divergences of opinion, resulting in order-of-magnitude variations in calculated distances for virtually identical design input data. Furthermore, operating experience has shown that a rigorous application of theory results in unnecessarily large distances. This has caused many compromises within the telecommunications industry. The most noted one is a newer standard, IEEE Std. 1590-2003, that lists a 150-m (~500-ft) mark as a default distance, if a GPR study has not been conducted at a given location.

24.8 ACKNOWLEDGEMENT

The authors thank the E&S Engineering staff, especially Elizabeth Smith and Chris Clemmens for their support and input. We thank Joe Joseph for his expert guidance and thoughtful comments and to Wayne Beaty at McGraw Hill for providing guidelines and editorial oversight. A special thanks to Robert D. Southy of Safe Engineering Services & Technologies Ltd. www.sestech.com, for his technical guidance and valuable insights.

24.9 BIBLIOGRAPHY

ANSI/IEEE Std 81, *IEEE Guide for Measuring Earth Resistivity, Ground Impedance, and Earth Surface Potentials of a Ground System*, 1983.

ANSI/IEEE Std 142, *Green Book—Grounding of Industrial and Commercial Power Systems*, 1982.

Arthur S., H. W. Beaty and H. Mahrous, *Handbook of Electrical Power Calculations*, 2d ed., 1997.

Donald G. F., and H. W. Beaty, *Standard Handbook for Electrical Engineers*, 14th ed., 2000.

Gilbert S., ABC of the Telephone, vol. 13—*Grounding and Bonding*, Revised 2d ed., 1999.

IEEE Std. 80-2000, *IEEE Guide for Safety in AC Substation Grounding*, 2000.

IEEE Std. 367-1996, *IEEE Recommended Practice for Determining the Electric Power Station Ground Potential Rise and Induced Voltage From a Power Fault*, 1996.

IEEE Std. 487-2000, *IEEE Recommended Practice for the Protection of Wireline Communication Facilities Serving Electric Supply Locations*, 2000.

IEEE Std. 1590-2003, *IEEE Recommended Practice for the Electric Protection of Optical Fiber Communication Facilities Serving, or Connected to, Electrical Supply Locations*, 2004.

Internet Grounding Reference: www.sestech.com

NFPA 70, *National Electrical Code*, 2005 ed.

SECTION 25
COMPUTER APPLICATIONS IN THE ELECTRIC POWER INDUSTRY

Tom Qi Zhang*
Senior Software Consultant, AREVA T&D Corporation

CONTENTS

25.1 INTRODUCTION

25.1.1 Growth of Computer Applications

The power industry is engaged in the generation, transmission, and distribution of electrical energy which is obtained by conversion from other forms of energy such as coal, gas, oil, nuclear, water, or other renewable energy. These activities often include mining, rail transport, shipping, slurry pipelines, and storage of energy in many forms. Many electric utilities are also engaged in the transmission and distribution of gas.

In the first 90 years of its history, the industry expanded at a pace nearly twice that of the overall economy, doubling roughly every 10 years. During this period, real prices per kilowatthour

*The author acknowledges the contributions of past authors and reviewers including James V. Mitsche (PTI), M. M. Adibi (IRD), and J. D. Cypert (IBM).

decreased steadily because of generation, transmission, distribution, technical improvements, productivity increases, and stable fuel prices. Throughout the 1970s, increased fuel costs, limits in economies of scale, diminishing returns in technology improvement, and increased regulation costs led to increased kilowatthour costs and reduced demand growth.

The political and economic response to increasing costs has been a movement to smaller generator sizes, minimization of capital investment, and attempts to control costs by fostering competition in generation supply. Incentives were also established to reduce demands and increase load factors. Today power supply is diversifying away from large central station technologies and toward increased use and availability of the transmission system.

In scheduling its day-to-day operation, and in planning for its future growth, the industry has made extensive use of analytical tools and mathematical models which, through optimization and simulation, help in the decision-making process. As a consequence, the industry has long been one of the largest users of computers and among the most sophisticated in its modeling and computational techniques. This use is quite understandable when one considers the high cost of power system equipment, the complexity of power systems, and the severe operational, reliability, and environmental requirements on the electricity supply.

Computer applications have assisted the industry in achieving its objectives: reducing the cost of energy delivered to consumers, improving the quality of service, enhancing the quality of the environment, and extending the life of existing equipment. These objectives have been achieved as follows:

1. Since the industry is one in which capital investment is usually high (over 10% of total spending by the nation's industries), unit costs have been reduced by operating facilities closer to their design limits, allowing better utilization of equipment.
2. Unit costs also have been reduced by automation, allowing operation with fewer personnel, and by optimization, lowering fuel consumption per kilowatthour delivered.
3. Electricity cannot readily be stored; therefore, production and consumption must be simultaneous. Hence enough capacity is required to meet the maximum coincident demand or peak load of all customers. Interconnections between power systems provide important economies arising from different time patterns or diversity of use of the component systems in the network. They allow higher power system reliability at lower capital cost.
4. Quality of service has been improved by reducing the number, extent, and duration of service interruptions, thus providing a more reliable service.
5. Quality of environment has been maintained by operating facilities within acceptable bounds of emission, thermal discharge, waste disposal, and more effective land use.

Today the industry has reached a stage where computer systems are no longer merely an engineering tool. The effectiveness of computer applications is one of the key elements in achieving the basic functions associated with the planning, designing, construction, operation, and maintenance of the power system. In fact, engineering and computers have been integrated. This integration may be viewed as tending toward the construction of a utility industry information system. Such a system is shown in Fig. 25-1. It depicts a typical information system which may be viewed as a combination and integration of several functional information systems.

Such an information system can extend the company capabilities by making relevant and current information accessible to both technical and management personnel. Designs can be refined by using measured data or operations experience, projects can be monitored, revenue requirements can be predicted more closely, and the experience of operations can be reflected in the methods and criteria used in planning and engineering. The information system thus can provide meaningful data at proper times and locations to make decisions and concentrate resources in the most effective manner.

Computers and their applications are ubiquitous in electric utilities. As in most industries, the business and corporate uses are extensive. This section deals with the sophisticated engineering and operations applications of computers, which are often unique and specialized to the industry's goals and technical demands.

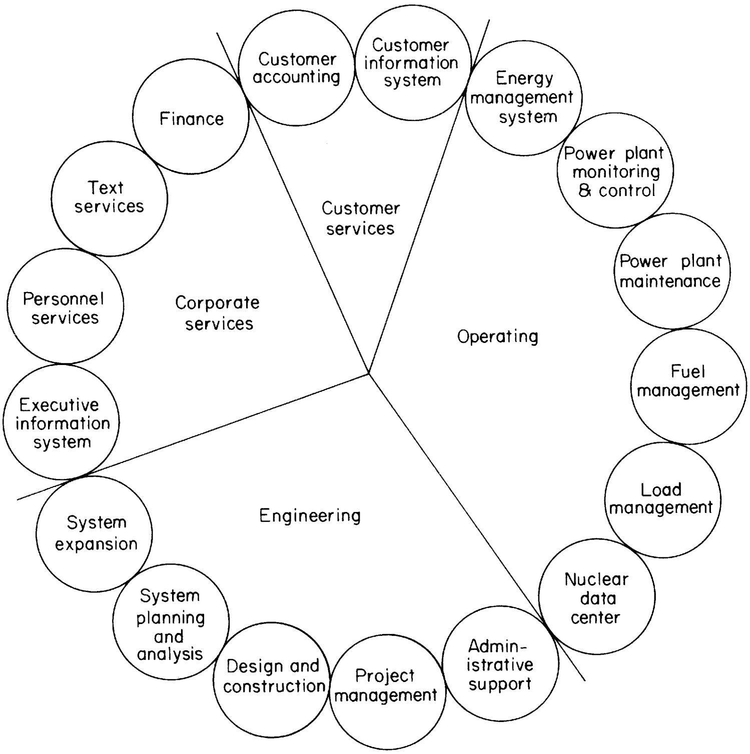


FIGURE 25-1 Electric utility information systems.

25.1.2 Goals of the Power Industry

The industry's purpose is to provide adequate, reliable, environmentally compatible electricity at reasonable cost with the ultimate goal of improving its productivity and net earnings. In spite of the differences between publicly and privately owned utilities, this goal is applicable to each, in different form. This goal is reached by pursuing a number of objectives as described below.

Improved Financial Management

- *Raising new capital*. Traditional electric utility companies and independent power producers are major utilizers of capital to finance and build new capacity, replace or renovate old equipment, and retrofit plants and delivery equipment for environmental and reliability considerations. Projected industry construction in the next decade runs into hundreds of billions of dollars. Competing demands for capital and its high cost encourage and justify precise planning, design, and operations.
- *Plant investment*. Utilities must spend very large sums in generating plants and transmission facilities. Present-day decisions on such additions, together with the proper selection of plant sites and the acquisition of transmission rights of way, have long-range financial implications affecting earnings. At present, the industry is experiencing difficulties in selection of plant sites and obtaining rights of way, licenses, and permits, with the results that the industry seldom obtains new plant sites and is forced to expand existing generating sites. Demand-side options must be properly weighed against generation expansion alternatives. Independent power producers (IPPs) must

make precise investments, and utilities must properly invest in facilities needed to utilize and accommodate IPPs. This situation compounds the problems of system modeling and system losses, and increases transmission system dependency.

- *Long-term contracts*. Fuel constitutes about 35% of the industry's total annual operating expenditures. A typical modern power plant consumes about 500 tons of coal each hour, and its average life is about 30 years. A nuclear power plant of a similar size requires an initial nuclear core costing hundreds of millions of dollars plus a significant annual refueling expenditure for the next 30 years. Independent power producers supply energy under contract for varied periods and conditions. The goal is to procure energy supply and these fossil and fissile fuels through long-term contracts providing a continuous supply of fuel at reasonable cost throughout the plant's 40- to 60-year life.
- *Growth through affiliation*. There have been a significant number of corporate mergers between large and small utilities. The goal in these affiliations is to meet the growth in demand for energy by taking advantage of economy of scale; consolidation of administration, engineering, construction, research, and development; and increasing reliability of bulk power supply.
- *Economy and reliability*. The industry has achieved significant improvement in economy of operations and in reliability of power systems either through direct operational pool functions or with contractual economical agreements.

Increased Revenue. For 30 years (1935 to 1965) utilities were, by lowering costs, reducing their rates and increasing sales. During this period of falling rates, owing to lags in regulatory rate adjustment, utilities enjoyed a higher revenue and were motivated to be efficient. The costs were reduced by the installation of larger generators, higher transmission voltages, lower fuel costs, and shifts to available gas and oil from coal.

From 1973 to 1990, the utilities went through a period of rising costs due to rises in fuel cost, environmental regulation, diminishing efficiency returns in technology (unit sizes and improvements), and investing in new technologies such as nuclear plants. During this rising-cost period, the regulatory lag in rate adjustments had an adverse economic impact, demanding detailed analysis of past and present operations and projection of future requirements by financial modeling, optimization schemes, and simulation techniques.

The expanded list of supply- and demand-side options and the desire to open transmission system access have made electric utility planning and operations much more complex. Construction delays and the desire to minimize capital expenditures have resulted in the electric power system being used in unexpected ways, and design safety margins must be reduced or stressed.

Reduced Cost. Cost reduction can be achieved by reducing investment per kilowatt of installation capacity for generation, transmission, and distribution and reducing operating cost per kilowatthour of energy delivered.

Reduced plant investment can be achieved by proper generation mix and location, increased transmission and distribution voltages, power pooling, interconnection planning, and coordination to gain further advantages of scale. Involved also are improved production and distribution facility utilization (i.e., capacity factor and load factor) through peak shaving, reserve sharing, load diversity, and distribution load balancing. Other means of reducing costs include designing facilities with more precision and reducing the factor of safety, reducing construction and inventory costs, and operating the facilities closer to their design limits.

Reduced operating expenditure can be achieved by adopting new technology that requires lower fuel costs, by improving conventional and established methods of higher energy conversion efficiencies, by reducing energy losses in transmission and distribution facilities, and by interchanging energy with more economical resources and different time zones in different seasons to take advantage of diversity.

Other means are producing and distributing electricity with fewer personnel; minimizing the labor force and material inventory required for maintenance, repairs, and restoration of generation, transmission, and distribution facilities; and reducing customer accounting, general accounting, and administrative expenses.

Improved Quality of Service. Among the requirements in this category are reducing the frequency, duration, and extent of outages in the power supply; reducing voltage and frequency discontinuities and sudden excursions (power-line disturbances) to sensitive electronic loads and digital equipment; and improving customer services through prompt response to inquiries, requests, or complaints. It is also important to maintain the power supply within prescribed ranges and specifications and to restore interruptions in service quickly.

Enhanced Environment. Means of improving environmental impact include reducing thermal discharge to natural bodies of water through the use of artificial lakes, cooling towers, and desalinization processes and advancing direct conversion of heat energy to electrical energy as by magnetohydrodynamics, thermionics, and fuel cells. Also involved in conventional systems are reducing the release of combustion products (sulfur dioxide, nitrogen oxides, carbon dioxide, and particulate matter) in the atmosphere; reducing the frequency, duration, and intensity of pollution concentrates in urban areas; and providing more productive uses for fly ash. Safer storage of nuclear waste is of primary importance.

In the design of systems, selecting remote or underground sites for generating stations, improving aesthetics by the increased use of underground distribution facilities, and beautifying transmission towers and lines in harmony with the countryside are all being urged by environmentalists. Modern transmission- and distribution-line designs reduce magnetic and electric fields in consideration of possible health effects.

Improved Employee Skill

- *Labor.* In earlier years, the power industry had a labor force of about half a million employees, a small force when compared with its very high output. In the 1970s, while the generating capacity doubled, the number of employees remained substantially the same. This was achieved through the operation of larger installations with fewer personnel, centralized control of generation and transmission, unattended substations, and minimizing maintenance and repair crews by automating dispatch procedures. This trend no longer holds.
- *Professional.* The design and construction of large installations such as generating stations and extra-high-voltage lines are often contracted out and are engineered and supervised by consulting firms. Thus, in effect, the consultants provide a common professional pool for all utilities. The electrical manufacturers have been primarily responsible for research and development of the industry, and the practice of accepting turnkey contracts is common. Thus manufacturers also provide a common pool of labor.

However, the advent of nuclear power, extra-high-voltage, and environmental limitations requires significant changes in utility systems and calls for an increase in both the quality and quantity of professional labor. The industry recognizes the need for this rapid increase in in-house skill. This can be provided by (1) improving the productivity and effectiveness of employees, (2) merging and affiliating with neighboring companies forming regional groups, and (3) maintaining aggressive in-house research and development as well as supporting institutions of higher learning and research organizations by sponsoring research and development efforts.

25.1.3 Spectrum of Computer Usage

A review of engineering and operating computer applications indicates that they fall within several broad categories, as shown in Fig. 25-2 and as described below.

System Expansion. These applications are related to 20-, 10-, and 5-year construction programs and cover planning, design, and construction of new facilities. These functions are performed at least once a year and use long-range load forecasts and other predictions as input data. Competitive pressures and complexity of expansion options demand that engineers have sophisticated interactive computer tools, decision-support and communication systems, and report-generating mechanisms.

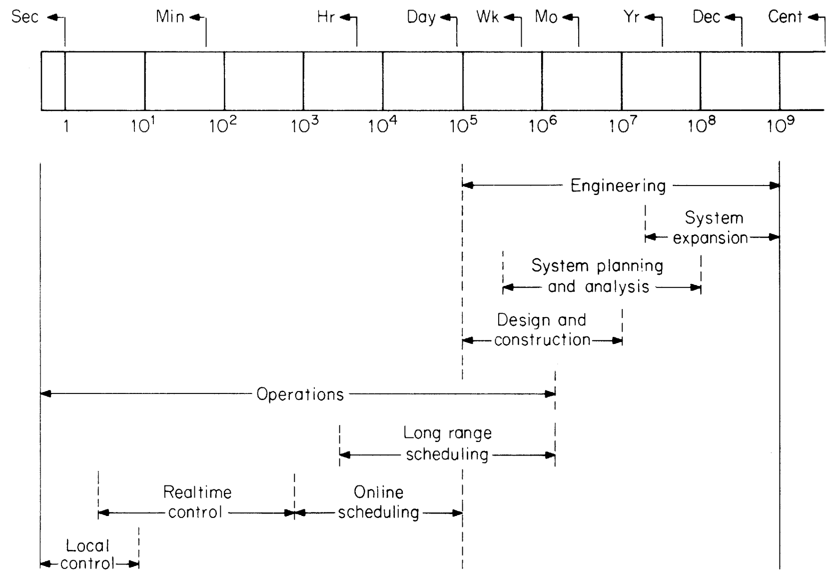


FIGURE 25-2 Spectrum of computer use.

System Planning and Analysis. These applications deal with 3- and 1-year construction of new facilities and the economic and reliable operation of these additions in conjunction with other interconnected power systems. Nuclear fuel management, annual hydrothermal coordination, and coordination for firm transmission and generation planning are among these functions. Because these programs are more frequently called on, they normally reside on disk storage devices. Thus only changes in data and programs need be entered when using specific programs.

Long-Range Scheduling (Operating). These applications are related to annual, monthly, and daily operation of the power system. In this category are transmission and generation maintenance scheduling, unit commitment and withdrawal, and other functions dealing with both reliability and economy of operation. Electric power systems are more complex and stressed than ever before. Maintenance of reliability and cost reduction require fast interactive computation in order to evaluate contingencies, operating options, and limits.

On-Line Scheduling. These applications are related to security monitoring and determination of reserve indexes and hourly data recording. These schedules are performed at least once an hour, although some applications such as pumped storage scheduling are performed weekly and daily. They are based on historical data but also need current power system data such as facilities in and out of operation, generation outputs, and line flows. Therefore, they require direct data flow into the computer. The results, however, are presented to the user for consideration and execution. Because of the scheduling nature of these applications, very fast data acquisition is not a prerequisite; however, accuracy and timeliness of schedules are related to the extent that they include direct data acquisition.

Real-Time Control. These regulating functions are carried out to meet the changing demands on the power system. Power system monitoring, security assessment, and display, rescheduling, and control of system frequency, tie-line flows, voltage conditions, and transmission flows are examples of this category. Other examples are closed-loop automatic control of generating units and interchange scheduling with neighboring companies and pool areas. These functions are performed in a time range of a few seconds to several minutes and therefore not only require direct data flows into the central computer but, in addition, require signals from the computer to the various remote controllers and actuators.

Local Control. These applications require a response speed beyond the capability of central computer control and related communication. Most of these functions are initiated immediately after a fault develops or a variable exceeds certain limits. Their objective is to react quickly and correct the situation or to isolate and contain a disturbance. These functions are performed in the few milliseconds to several seconds range and can best be handled by local computers: (1) by directly sensing variables and controlling through actuators (e.g., direct digital control of boilers or digital relaying of the substations) and (2) by superimposing the computer on the local controllers or protective relays in order to reset their operating positions. The latter applications are in the 1- to 10-s range.

The computational requirements shown in Fig. 25-2 cover both engineering and operating functions. These areas of computer activities are interrelated. From the preceding discussion it is clear that the power system operating functions do not have to be performed necessarily in real time.

25.2 ENGINEERING APPLICATIONS

As the electric utility industry has grown in size and complexity, modifications and additions to existing electric power networks have become increasingly costly. Therefore, it is vital that different design possibilities for additions and modifications to the network be studied in detail to determine their effect on the network, their effect during abnormal operating conditions, and their applicability as a flexible solution to current and future power demands.

The design and construction of planned facilities involves the efforts of a sizable engineering staff and a substantial investment in facilities. To provide support in these activities, computer programs have been developed for analysis of specified designs. The application of these programs contributes to the installation of reliable and economic facilities. The major engineering applications are shown in Fig. 25-1 and summarized below.

25.2.1 System Expansion

The system-expansion applications (Fig. 25-3) support the long-term (5 to 20 years) planning function for generation and transmission of power. The system-expansion application area represents the typical decision support environment in that many cases are produced and a variety of options and strategies are considered in the planning process. This area controls large common data sets from multiple sources. Lengthy reports are produced for internal documentation and regulation approval. In the past, most of the processing was batch-oriented because of the length of computation. Today on-line dialog with the applications is feasible and essential for evaluation of alternatives.

With the current economic outlook, the majority of emphasis in the industry will be to develop more efficient use of existing facilities rather than new construction. Load forecasting and production costing are becoming the most significant items in system expansion to predict load requirements and operating costs. Tradeoffs between expansion and new facilities are increasingly important. The applications in this category are as follows:

Load Forecasting. This application is the basis for all planning functions. It utilizes historical data, trends, economic factors, and residential and industrial projections by geographic area to produce load requirements and load duration plots by area. The effects of demand-side management are incorporated to evaluate the most cost-effective options. It also predicts the load factor.

Generation Mix Analysis. This plans the optimal mix of peaking, base-loaded, or independent power producer units; fuel type; and location of units to meet the future load requirements. It also provides a buy-and-sell analysis and accounts for reliability of generation.

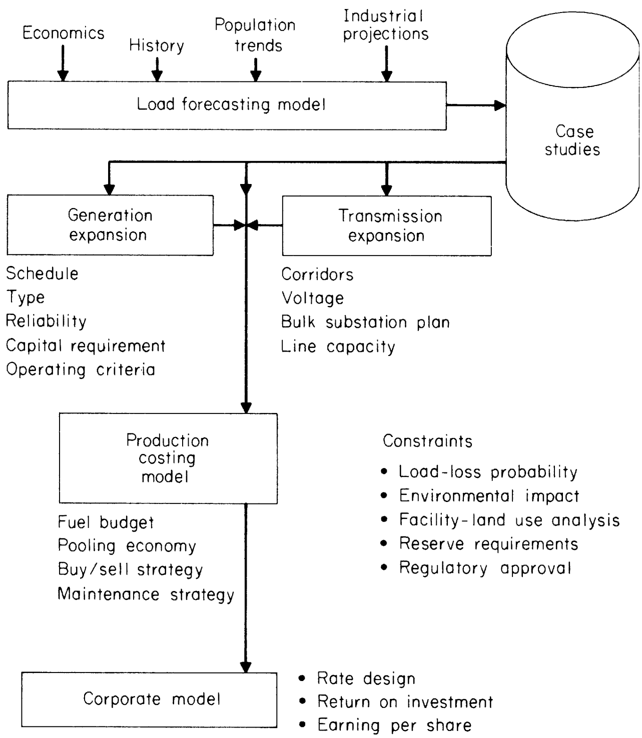


FIGURE 25-3 System expansion applications.

Production Costing. This stimulates the operation of the existing and planned generation facilities for several years in order to predict the fuel budget. It meets the load forecast and accounts for the generation availability and the hour-by-hour dispatch of generation. New techniques use statistical approaches versus detailed models.

Loss of Load Probability. This accounts for unit availability and the reserve requirement to produce a probability of loss of specific loads.

Voltage Level Analysis. This application is a tool to plan voltage levels of existing and planned transmission facilities. It provides tradeoffs of network losses versus capital requirements.

Environmental and Facility Land-Use Analysis. This set of applications assists the planner in locating plants, substations, transmission towers, and lines. Tradeoffs considered are expansion versus new facilities, right-of-way utilization, and environmental impact of planned facilities. Tighter environmental controls are increasingly affecting expansion and operating plans.

25.2.2 System Planning and Analysis

This application area supports the short-term planning process and provides tools for analyzing incremental expansion. System planning and analysis applications are high in floating-point content and represent a significant computational requirement. Many cases are analyzed, and there is a rapid turnaround requirement. While there is on-line dialog with the application and it is common to

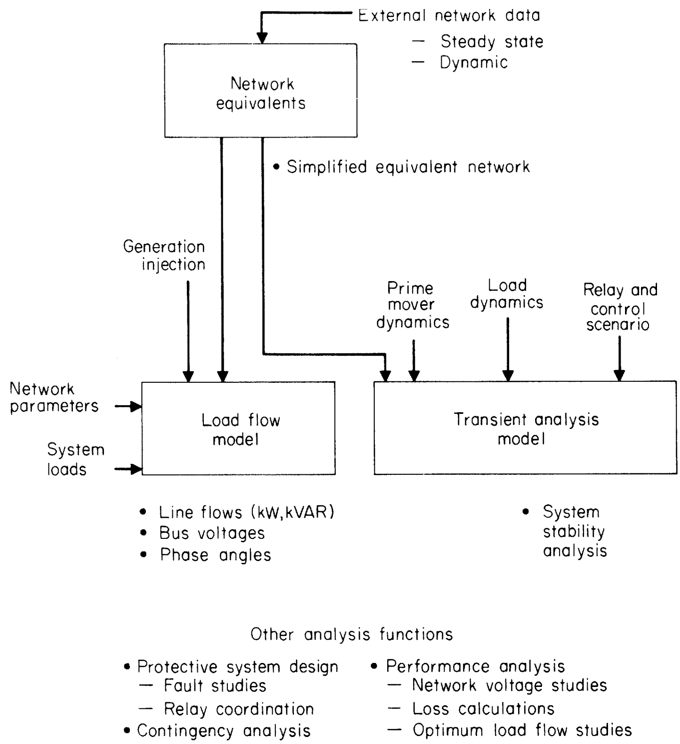


FIGURE 25-4 System planning and analysis applications.

provide on-line display and edit of results, the trend is to use interactive graphics in all phases of this decision support process. The specific applications are shown in Fig. 25-4 and described below.

Load Flow. The load-flow program is one of the major tools of system planning and is utilized extensively. Important to the system planner are that input data errors be minimized and that there be an easy and rapid turnaround for answers when the frequency of program use is high. To accomplish this, interactive capability is provided with the ability to store base cases or numerous power system models on the computer's disks. The storage capability provides many different cases that the planning engineer can access for studying or varying a particular system condition. Load flow enables the power system planning engineer to simulate and solve various power system expansion alternatives in an interactive mode. It utilizes a graphics color terminal specified with a special set of graphic characters that presents results in the form of system on-line diagrams. The multicolor feature of the terminal is used to indicate heavily loaded lines, bus voltages outside normal limits, and open circuit breakers. The engineer working at such a terminal may, with a mouse and alphanumeric keyboard, remove, add, or change elements of the system being studied and request a solution from the host computer.

The simulation programs associated with the load flow program are a system of linked programs that have the following capabilities:

1. Basic programs involving calculation of voltages, power flows, angles, and interchanges between areas of a power system.
2. A network reduction program to represent large networks as equivalents in conjunction with the specified area to be studied.

3. A distribution-factors program that indicates the sensitivity of response of the various circuits to outages of specified transmission lines, used to predict thermal limits with linear (superposition) techniques.
4. The series of programs associated with the stored load flow files which permit accessing a particular case, deleting a case, adding a case, and changing a case.
5. The var allocation program which selects the minimum amount of kilovars of compensation necessary to maintain bus voltages within specified limits under normal and/or emergency conditions. Optimal power flow (OPF) is a versatile alternative for var planning, economic dispatch, and performance improvement.

Transient Analysis. This is a large, dynamic simulation of the generation and transmission network in the transient state. It models the dynamics of synchronous machines after a system fault condition. It produces network stability information that guides the engineer in design of the network and its protective system.

Protective System Design. This application automates the work of the relay engineer in designing the protective system. It includes fault and relay-coordination studies and calculates complex relay settings.

Switching Surge Analysis. This program calculates the voltage and current transients resulting from switching surges and lightning strikes.

Contingency Analysis. This is the off-line analysis of predetermined outages and network contingencies. It is used by planning in design studies. Results are used as input to the operating area for guidance during problem or alert conditions. When combined with OPF, this is referred to as *security-constrained optimization* (SCO).

Performance Analysis. This attempts to satisfy voltage level criteria for the network and specific consumers while minimizing network losses.

25.2.3 Design and Construction

This function includes all routine applications associated with the design and construction of power plants, transmission facilities, substations, service centers, and distribution facilities. Electric utility companies on average spend more for new construction each year than any other industry. Although their primary objective is the usual production and sale of a product, they must be concerned with a large capital investment program. Efficient planning, scheduling, and control of labor and material resources are necessary if customer demand is to be met and at the same time a fair rate of return is to be provided on the investment.

Major Programs. Design and construction is a multidiscipline activity that employs computer applications dealing with electrical, mechanical, and civil engineering functions. A partial listing of specific applications is described below.

The *tower analysis* program provides a summary of the maximum tension and maximum compression for each member of a three-dimensional structure over the entire load range specified. This program also spots structure locations, plots a profile of the transmission line, and calculates sag, insulator swing, and ground clearance.

Line sag calculates sag and tension of conductors under a given situation.

Branch circuit design uses the load, distance, number of cables in a raceway, wire temperature rating, motor-starting, and full-load amps to compute the voltage drops and sizes of breakers, cable, and conduit in a circuit.

Structural design programs are used to design concrete and steel structures using as input the structure configuration and loads such as floor, roof, and impact.

The *structural steel framing* program is used to design the beams, columns, girders, and base-plates of power plant structures.

The *foundation-slab analysis* program is used to design large, complex foundation mats. The results permit evaluation of various slab thicknesses, soil bearing pressures, shears and bending moments, and reinforcement areas.

The *concrete stack analysis* program is used to analyze proposed stacks by determining loadings, resulting stresses, and required steel reinforcement. This program is used extensively in the design of very tall concrete stacks selected for new power plants.

Piping programs are used to perform stress analysis of piping systems and determine hanger design information. The *power plant piping* program analyzes the flexibility of a piping system under the influence of temperature.

The *cable routing* program provides the shortest cable route between nodes, percent raceway fill, and number of cables in a tray or raceway.

Interference analysis resolves the interference between pipe, cable tray, and structures occupying the same space.

The *heating, ventilating, and air-conditioning design* program uses thermal loads and the building configuration to calculate the size of refrigeration equipment and ductwork required.

Fluid dynamics analysis analyzes piping systems for pressure drop, flow distribution, and power requirements.

Hydrologic analysis is used to determine seepage flow networks, underground flow, and rainfall and runoff drainage for culvert and bridge size and design.

Earthwork design is used to design embankments and roadways and perform settlement and embankment stability analysis.

Geotechnical evaluation is used to evaluate soil testing results and determine the strength and swell of soil for dam and foundation design.

Foundation design programs are used to design foundation pile, pier, mat, and spread footing.

A *statistical analysis of equipment failures* means data collected on the frequency and cause of equipment failure is used to determine the likelihood of similar failures in the future based on various changes. The purpose of establishing an equipment operation database is to record and summarize the specific causes of service interruption to generation, transmission, distribution, and communication system equipment as well as to customers. The data provide the basis for designing new systems to specific reliability levels, monitoring equipment and manufacturer adherence to desired availability standards, and carrying out maintenance scheduling activities.

The database consists of a main file for each major equipment category and supplementary files which supply input to a family of programs designed to provide the engineers with periodic statistical reports. Engineers also have the ability to retrieve from this database any combination of data of their own choosing.

Transformer load management consists of a series of programs to process manufacturing performance data for distribution transformers and derive an economic evaluation based on unit cost and expected loss contributions over the expected lifetime. Since distribution transformers represent such a substantial proportion of system investment, it is imperative that they be utilized to their fullest economic capability.

A large percentage of distribution transformers are nominally underloaded; that is, they are oversized for the load being served and hence waste money through overinvestment and excessive core losses. Overloaded transformers also waste money in terms of copper losses, loss of life, fuse and transformer burnouts and replacements, and the investigation of low-voltage complaints.

Drafting includes engineering sketches and standard symbols used to lay out a drawing on a terminal and the results are printed or plotted. Included in this application are computer-aided design and drafting packages.

Economic analysis is used to make economic decisions between alternative sets of system designs or equipment.

Most of the preceding applications require common data. The trend is to treat the data as a corporate resource and to capture and maintain them in a common database. This provides for consistency of data, avoids duplication, and minimizes errors due to entry of the same data in different programs. This common database can then be used by many of the design and construction programs.

Resource-management subsystems that support the design and construction applications are briefly described below. The purpose of resource management is to help utilities in more effective utilization, control, and management of their basic resource (people, equipment, and facilities) in the distribution system.

Distribution Construction Information System. This application supports the management of new investment and maintenance in the distribution area. It provides information for planning, scheduling, and controlling equipment, labor, and material resources and becomes a tool to assist public utilities in providing consistent service and meeting customer demands while realizing fair rates of return on capital investment.

The term *distribution construction* refers to the entire process of work requesting, design, scheduling, reporting, and closing of that portion of the facilities closest in service to customers. New distribution work stems from three types of activity: system maintenance and improvement requirements, customer requests, and inspections or surveys.

For many utilities the distribution system alone represents close to one-third of the total capital investment. Thus it is not surprising to find continuing concern with the process by which facilities are constructed and maintained and by which costs are transferred to property accounts or charged to expense appropriations. This concern is often focused on improving the distribution work process to support the planning, design, scheduling, controlling, and tracking of jobs. Such improvements are undertaken to obtain a more effective and efficient work process leading to an earlier plant-in-service and to improving the utilization of the many resources devoted to distribution construction, maintenance, and operating tasks.

Computerizing the distribution work process is desirable because of the significant capital investment and expense components of utility costs and because the work process has characteristics that lend themselves to a high degree of computerization and to improved productivity and control.

Distribution Facilities Information System. This application provides the information required to plan, control, maintain, locate, account for, and manage the distribution facilities of an electric utility. It is also referred to as an *automated mapping and facilities management system*. When combined with terrain and other landmark information, it is referred to as a *geographic information system* (GIS). It is composed of a graphics system and a database system.

The graphics system provides the interactive functions required to capture information needed to maintain a database for facilities by locations. The user is required to define to the graphics system the facilities and the data elements which are associated with them. This includes its data fields, the pictures that are displayed on a map to represent it, and the connectivity requirements, if it is a network facility.

The graphics system employs a graphics workstation composed of at least one high-resolution display, an alphanumeric display, a keyboard, a mouse, and various hard-copy devices. This workstation is used to enter geographically related data, making it subsequently possible to display the data pictorially (maps), interact with the display (zoom in, window, edit, etc.), display facilities data and alter them, or to make additions or revisions. These are functions that formerly involved manual drafting and filing methods.

Maps or data generated at the workstation may be stored in a common facilities database accessible to many users. The manner in which these data are stored varies. Some systems are able to retrieve only the map facets that were entered; the user must establish the relationship between adjoining facets. In other systems the data exist as a continuous network, and the user requests only the portion and type of data needed. Storage techniques based on the common corporate database management system are becoming prevalent because of the common requirements for facilities data by many departments throughout the company.

The production of maps, diagrams, and pictures is a by-product of this system. Of far greater importance is the network relationship of the facilities data. This allows such applications as feeder analysis, transformer load management, fuse coordination, branch circuit design, and fault current calculations to be executed using the common facilities database.

Material-Management Information System. This system is used to plan for and control the flow of materials in and out of the company. A material items database may be accessed by many departments for multiple purposes. The main subsystems are stores operations, materials planning, and purchasing.

Stores operations relates to all day-to-day activities within the warehouse location. Included in these are functions such as stock inquiry handling, recording of stores transactions (receipts, issues), item location management, order and requisition initiation, and material reservation and allocation control.

Materials planning refers to the control and management of an inventory, both repairable and expendable parts. The functions under this application are acquisition analysis, item forecasting, materials requirements planning, reporting, and stock taking control.

The purchasing area includes the functions of ordering material from the suppliers and transferring to the inventory on receipt of the material. The functions within this are purchase order writing, quotation preparation, receiving, returns, implementation, quality assurance, vendor performance analysis, and invoice matching.

25.2.4 Project Management

The objective of project management is to control project costs and schedules in the maintenance and construction of power system facilities. Power engineers have used manual project information systems for years. Now there is rapid movement to automation of project control with provisions for on-line display and edit of results.

Project management in the engineering departments includes project control, project scheduling, and resource optimization tools. These tools are required to manage small procurement projects. However, they also could be major, long-term projects such as construction of a facility, installation of a major program, or daily tracking of activities within a department. Automated techniques, taking into consideration the control of time, resources, and costs, allow more productive utilization of project management personnel and stricter control of projects than manual methods. There are three major components of a project management system:

1. *Critical-path method*. A network represents a project which consists of a mixture of serial and parallel activities and employs a combination of personnel resources, materials, and facilities. When time is associated with each activity within a network, critical-path methodology can be used to analyze the network and determine the longest time path to completion of the project. All other time paths through the network will then have some slack in terms of the critical path.
2. *Resource management*. Project management and scheduling provides the means to plan and control a variety of projects. These systems permit tasks to be scheduled, resources assigned, costs allocated, and progress reported. Using this process, management can address identified problem areas and adjust its plans accordingly.
3. *Project costing and estimating*. Cost control techniques involve the ability to estimate and assign costs for labor, material, facilities, test equipment, and other resources to all activities comprising the execution of all phases of a project. In addition, some application programs permit extending rates; accommodate matrix and other organization structures; compute general, administrative, and overhead expenses; and summarize project costs over selected parts of projects as well as multiproject groups.

25.2.5 Administrative Support

Administrative functions have been automated to serve the various requirements of engineering departments. Because these requirements are common throughout the company, integrated or common solutions are often used. Administrative support typically falls into the following categories.

Text processing is the preparation, output, and data entry and editing of text using an interactive host-based or stand-alone computer system. This service may be used by a secretary to compose a letter or modify an existing memo. It may also be used directly by an engineer for notes, lists, progress reports, and general documentation. Text processing in a power company is used to prepare and maintain operating standards and procedures, standard material lists, nuclear records, training manuals, maintenance and safety procedures, regulatory reports, and specifications.

Administrative processing allows the user to manage electronic document images. It includes activities such as copying and reproducing, document distribution, records file processing, mailing, and office correspondence. Documents may be filed on disk, searched for, and retrieved. Document search may be by name or by complex search parameters as in nuclear records.

Other administrative services provide a convenient means for writing notes, reminders, messages, and appointment records. Typical functions include calendaring, tickler file (diary), meeting schedule, phone list, and to-do lists.

Text and data integration applications provide the ability to include data created outside of text applications to form reports and letters. These may be used for the creation of manuals which include specification data to reduce redundant keystrokes and increase accuracy.

Communications applications provide an informal and unstructured method of communication within an organization. This provides the means of handling messages that might otherwise require a phone call or memo. It also allows for distribution to multiple locations and receipt acknowledgment.

Personal computing provides engineers with the tools, packages, and techniques that allow them to enter and manipulate data and accomplish in hours what might otherwise take weeks. The intelligent work station provides local processing, user-friendly interfaces, and access to host applications. Many times, all applications programs can be executed on a personal computer.

Presentation graphics applications are used to present data in a pictorial form. The graphics may be displayed on a terminal or converted to a hard copy using a printer, plotter, terminal copier, or an attached camera device. Typical uses include the presentation of engineering or statistical data as line, bar, or pie charts, preparation of foils for a presentation, or drawing sketches or diagrams for inclusion in a publication.

25.2.6 Power Market Computer Simulation

Computer simulation has been a powerful tool for electric power system researchers and operators to study the system behavior under various conditions. The increasing complexity of power systems and interconnections to other infrastructures, vulnerabilities to cascading failures, interactive and large-scale nature of these networks, coupled with advances in modeling, computational methods, software technologies, simulations, control of networks, and economic aspects, have driven the further development of power system computer simulation tools. With the advent of deregulation, unbundling, and competition in the electric power industry, new software tools are needed to improve the efficiency of that network without seriously diminishing its reliability.

Power market simulator is a software tool that helps clients simulate hedging strategies in electricity markets before the strategies are put into practice and account for market contingencies in market operations and production. Capable of modeling complex power market interactions, this software tool will also allow market participants to train their staff to effectively address market scenarios that, until now, could not be handled confidently. The market simulator is useful in helping to perform a variety of functions:

- System contingency analysis to develop hedging strategies for system contingencies
- Market contingency analysis to develop hedging strategies for market contingencies, such as changes in price caps, bidding mechanisms, and new entrants
- Congestion management to develop hedging strategies for congestion charges
- Planning and expansion to demonstrate improved efficiency, reliability, and service from planned generation expansion in specific network locations
- Interregional coordination to simulate seams issues and help test strategies for coordinating production across multiple markets

25.3 OPERATING APPLICATIONS

The prime concern of the electric utility industry is to meet the consumer's power demand at all times and under all conditions. Electric utilities are continually seeking and have been most receptive to every available technique which would reduce the capital investment per kilowatt of installation, reduce the operating expenditures per kilowatthour of energy delivered, and improve the quality of service to the consumer.

Computer systems can be found in use at all levels of operation in power systems. At the generating-plant level, for example, they are used to control and monitor unit startup and operating conditions. In bulk power substations they serve such functions as monitoring, event recording, and switching. And at the system operating center level, they help to improve the economy of operation, improve the quality of service, and simplify system operation. The major computer applications for the operating area of a utility are shown in Fig. 25-1 and briefly described below.

25.3.1 Supervisory Control and Data Acquisition System

The electric utility SCADA (Supervisory Control and Data Acquisition) system is now a mainstay. Most electric utilities have means to monitor their power system activity and control substation equipment from a central location that would be classified as a SCADA system. The long sought improvements in efficiency promised by upgrading manned substations to monitored substations have been largely achieved.

It has taken many decades to get to this point. There has been a long evolution of technology and change that has brought SCADA technology this far. While substation automation is considered current technology, it is valuable to understand the steps in technological evolution and to recognize that some of that history is still in use in utilities today. The evolution in computer technology that helped drive the SCADA systems in their infancy has seen many significant changes. In the substation, microprocessor technology that gave birth to the intelligent electronic device (IED), opened the opportunity to combine functions and has significantly changed the landscape for the substation interface.

The SCADA system connects two distinctly different environments. The substation, where it measures, monitors, and digitizes; and the operations center, where it collects, stores, displays, and processes substation data. A communications pathway connects the two environments. Interfaces to substation equipment and conversations and communications resources complete the system. The substation terminus for traditional SCADA system is the *remote terminal unit* (RTU) where the communications and substation interface interconnect.

SCADA system RTUs collect measurements of power system parameters and transport them to an operations center where the SCADA master presents them to system operators. Predominantly, there are real and reactive power flow (watts and vars) voltages and currents. But other measurements like tank levels—pressures and tap positions are common to SCADA systems. These belong to the class of measurements termed analogs. Almost anything that can be viewed as a continuous variable over a range fits this category. Analog data is refreshed periodically so that the operator can be assured that data on this screen is relevant. The refresh rate is often dependent on the characteristics of the data being viewed and the communications resources available.

SCADA master stations monitor the incoming stream of analog variables and flag values that are outside prescribed limits with warnings and alarms to alert the system operator to potential problems. Data are screened for "bad" (i.e., out of reasonability limits) data as well.

SCADA systems also collect the state of power equipment such as circuit breakers and switches. These data are presented to the system operator, usually on graphical displays, to give the operator a view of the connectivity of the power system at any given moment. Various state change-reporting techniques have been used to report such changes for the system operator. These include flagging momentary changes, counting changes and time tagging them with varying degrees of resolution (sometimes as short as one millisecond).

SCADA systems almost always provide a means for the system operator to control power equipment. This includes circuit breakers, switches, tap changers, and generators. It may include some peripheral equipment in the substation as well.

In the operation center, a SCADA system has at least one computer, communicating to substations and/or generating stations collecting data, issuing control commands, and storing the incoming data. The system operator views data and messages through a set of displays on "view stations." The displays allow the operator to control power equipment and make system changes through a screen dialog.

Besides these basic functions, the operations center computer archives data and displays selected data sets, such as trends and logs in special ways for the operators. More modern systems provide data to other areas of the utility enterprise in any number of different forms and services.

25.3.2 Energy Management System (EMS)

New Development of EMS/SCADA Technology. As the central nervous system of the power network, the control center—along with its energy management system (EMS) (see Fig. 25.5)—is a critical component of the power system operational reliability picture. The market's shift, beginning in the late 1990s, from an investment strategy focused on improving capacity/reliability to one geared toward meeting the needs of a deregulating market impacted not only day-to-day infrastructure investment but also research and development. At the same time, utilities decreased EMS budgets even further because generation scheduling and optimization came to be viewed as a function of the deregulated side of the new market structure. Transmission entities therefore slashed their EMS budgets and staff, thinking that local independent system operators (ISOs) would assume

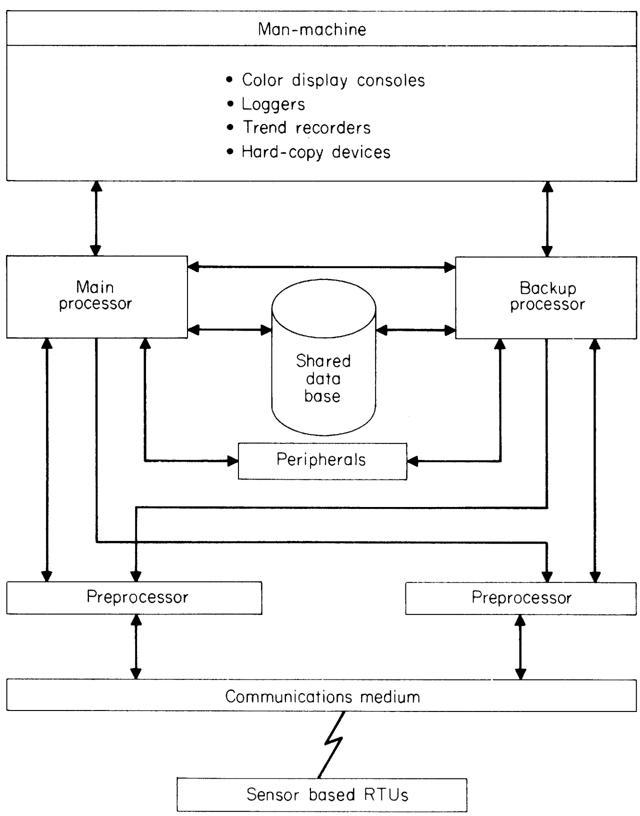


FIGURE 25-5 EMS architecture.

responsibility for grid security and operational planning. Those factors, combined with the fact that EMS technology has historically lagged behind the IT world in general, has created a situation where control room technology is further behind today than it has ever been.

The following important technology trends will come into play in the new EMS/SCADA solutions:

- *Visualization Control.* Control-room visualization today is still limited primarily to one-line diagrams, which are insufficient when it comes to today's needs to understand the availability of electricity at any given time and location and in understanding load, voltage levels, real and reactive power flow, phase angles, the impact of transmission-line loading relief (TLR) measures on existing and proposed transactions, and network overloads.

  Three-dimensional, geo-spatial, and other visualization software will become increasingly indispensable as electricity transactions continue to increase in number and complexity and as power data, historically relevant to a contained group of entities, is increasingly communicated more widely to the ISOs and RTOs charged with managing an open grid. Not only do visualization capabilities enable all parties to display much larger volumes of data as more readily understandable computer-generated images, but they also provide the ability to immediately comprehend rapidly changing situations and react almost instantaneously.

  Three-dimensional visualization is an invaluable tool for using abstract calculated values to graphically depict reactive power output, impacts of enforcing transmission line constraints, line loadings, and voltages magnitudes, making large volumes of data with complex relationships easily understood.

- *Advanced Metering Technology.* In this age of real-time information exchange, automated meter reading (AMR) has set new standards by which the energy market can more closely match energy supply and demand through more precise load forecasting and management, along with programs like demand-side management and time-of-use rate structures. Beyond AMR, however, a host of real-time energy management capabilities are now on the market, which, through wireless communication with commercial, residential, or industrial meters, enables utilities to read meters and collect load data as frequently as once every minute.

  This enables utilities to better cope with dynamic market changes through real-time access to the critical load forecasting and consumption information needed to optimize decision support. The convergence of demand-response technologies and real-time pricing, wireless communications, and the need for more reliable and timely settlement processes are all drivers for enhanced metering capabilities. This, in turn, will create a demand for EMS solutions capable of handling much larger volumes of data and the analytical tools to manage this data.

- *More Stringent Alarm Performance.* The 2003 blackout drew attention to what has become a potentially overwhelming problem—SCADA/EMS has little or no ability to suppress the bombardment of alarms that can overwhelm control room personnel during a rapidly escalating event. In a matter of minutes, thousands of warnings can flood the screens of dispatchers facing an outage situation, causing them to ignore the very system that's been put in place to help them.

  Although distribution SCADA has been able to take advantage of straightforward priority and filtering schemes to reduce the alarm overload, the transmission operations systems have not. This is because transmission systems are networked, and it is more difficult to analyze the alarms to determine what needs to be shown to help the operator reach a conclusion. Also, reaction time is not an issue in distribution, and there is more value in taking the time to locate the fault before taking action; short outages can be tolerated. Other industries, for example, telecom, networking, and refining have had good success with inference engines and other rule-based systems for diagnosing alarm conditions and providing operator assistance. These are worth a second look by the EMS fraternity today.

  New analytical tools are needed in the EMS to enable operators to manage and respond to abnormal events and conditions. See Fig. 25-6.

- *Data Warehousing.* For many years, utilities have been archiving the operational (real-time) and nonoperational (historic) information captured by energy management systems. Today's thought leadership shift is to focus on how this archived operational and nonoperational data can be combined with emerging analytic functionality to meet a host of business needs, for example, to more readily identify parts of the network that are at the greatest risk of potential failure. If integrated

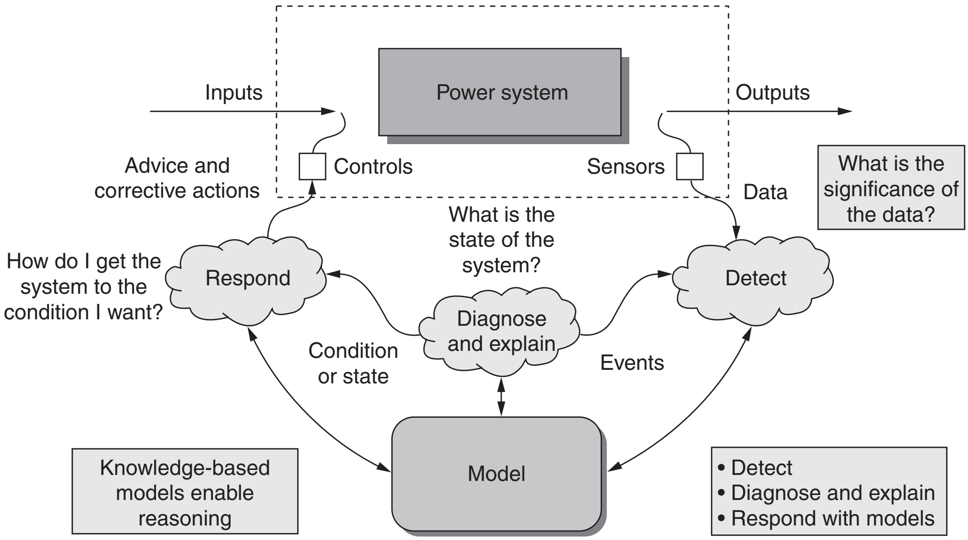


FIGURE 25-6 Real-time event management. (Courtesy of Gensym Corp.)

properly, heads-up information stored by these systems can also aid utilities in proactive replacement or reinforcement of weak links, thus reducing the probability of unplanned events.

A data mart is a repository of the measurement and event data recorded by automated systems. This data might be stored in an enterprise-wide database, data warehouse, or specialized database. In practice, the terms data mart and data warehouse are sometimes used interchangeably; however, a data mart tends to start from the analysis of user needs, while a data warehouse starts from an analysis of what data already exist and how it can be collected in such a way. Desktop access to the data each user needs to perform his or her job functions with an emphasis on business intelligence and knowledge management.

- *Communication Protocols.* EMS systems must have the capacity to talk to "legacy," i.e., preexisting, remote terminal units (RTUs) and, thus, are severely handicapped today in that many still rely on serial RTU protocols that evolved in an era of very limited bandwidth. As a result, most EMS solutions in use today are unable to exploit breakthroughs in communications, in particular, secure communications such as encryption and validation. This will need to change. Eventually, the need for encrypted, secure communications to the RTU, combined with adoption of substation automation and substation computers, may lead to the end of RTU protocols as we know them today and adoption of a common information model (CIM)-based data model for the acquisition of field data.
- *Enterprise Architectures.* To achieve the benefits offered by the technologies described here, EMS solutions need to be able to take advantage of modern enterprise architectures (EAs). EMS systems are typically not included as part of utility EA initiatives, but as their importance becomes readily apparent, this will change. Though EA definitions vary, they share the notion of a comprehensive blueprint for an organization's business processes and IT investments. The scope is the entire enterprise, including the control room, and, increasingly, the utility's partners, vendors, and customers.

  Though no industry-standard technical/technology reference model exists for defining an EA, it is clear that component based software standards, such as web services, as well as popular data-exchange standards, such as the extensible markup language (XML), are preferred, as are systems that are interoperable, scalable, and secure, such as Sun Microsystem's Java 2, Enterprise Edition (J2EE) platform, or Microsoft's.Net framework.

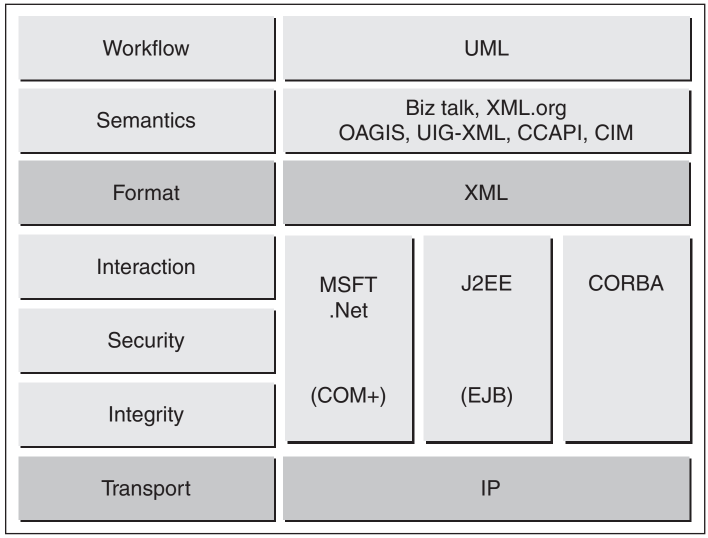


FIGURE 25-7 Integration standards.

It is also clear that frameworks and initiatives, such as the Zachman framework, Federal Enterprise Architecture Framework (FEAF), The Open Group Architecture Framework (TOGAF), and Rational Unified Process (RUP), will strongly impact how enterprise architectures for utility control operations are defined and implemented. See Fig. 25-7. By using shared, reusable business models (not just objects) on an enterprise wide scale, the EA provides tremendous benefits through the combination of improved organizational, operational, and technological effectiveness for the entire enterprise.

- *Web Services Architecture.* There are no EMS deployments today that take advantage of modern Web services architecture, although the architecture is providing tremendous benefits to businesses around the world and holds big promise for control room operations.

Like object-oriented design, Web services encompass fundamental concepts like encapsulation, message passing, dynamic binding, and service description and querying. With Web services architecture, everything is a "service," encapsulating behavior and providing the behavior through an API that can be invoked for use by other services on the network. Systems built with these principles are more likely to dominate the next generation of e-business systems, with flexibility being the overriding characteristic of their success.

As utilities move more of their existing IT applications to the Internet, a Web services architecture will enable them to take strong advantage of e-portals and to leverage standards, such as XML; Universal Description, Discovery, and Integration (UDDI); Simple Object Access Protocol (SOAP); Web Services Definition Language (WSDL); Web Services Flow Language (WSFL); J2EE; and Microsoft.NET.

The dispatcher's interface to the EMS system is through color display terminals used to display on-line diagrams and tabular displays. These displays are updated cyclically or whenever a change is detected. Various data and operating conditions are highlighted with color. The dispatcher uses the displays to interact with the system such as acknowledging an alarm, changing operating limits, controlling devices, determining system status, or executing a program.

The major functions performed by an EMS system are shown in Table 25-1. The types of processing that these functions include are cyclic work that must be completed within a specific cycle, event-driven work that results from random alarms and events that occur on the network, real-time or demand work that has specific response-time requirements, and interactive work which is dispatcher-initiated that also has response-time requirements.

TABLE 25-1 EMS Functions

Network surveillance and control	Generation commitment and control
• Network data acquisition • Data conversion, limit checking • Alarm processing • Logging and reporting • Load-shedding power device control	• Automatic generation control • Economic dispatch • Load forecasting • Unit commitment • Reserve monitor
Interchange management	**System security**
• Production costing • Transaction evaluation • Interchange scheduling • Interchange accounting	• Network status • State estimation • Contingency analysis • Interactive load flow • Training simulator

Network Surveillance and Control. This consists of continuous scanning of remote sensor-based units, acquiring all key network and generation data on a 2-s cycle, checking the data for problems, and presenting the status of the network to the dispatcher through displays. The control function allows the dispatcher to take selective control action on power system devices such as circuit breakers. Frequently these functions are called SCADA (supervisory control and data acquisition).

Network data acquisition programs acquire power system data through periodic scanning of local and remote sensors attached to RTUs. The raw data are checked for missing or invalid data before being sent to the central processors by the communications front end. Retransmissions are requested for missing status data.

Data conversion and limit checking programs are used to process the raw data. The data are scaled and converted to engineering units and stored in a database for use by application programs. Reasonability checks are made on the data and values not found to be within limits cause alarms.

Alarm processing programs generate alarms for scanned data and calculated results as changes of network status are discovered, limits are exceeded, or other invalid conditions are encountered.

Logging and reporting programs store selected data in historical log files, where they are available for analysis through displays and reports.

Power device control functions provide for the control of power system devices and for the placement or removal of device tags. Control actions made by the dispatcher to activate power devices are verified, then forwarded through the network to the proper destination.

Load shedding provides rapid access to data for guiding the action of the dispatcher during abnormal operating conditions. Typical of the data that are displayed is the amount of load relief that can be obtained through emergency output of all generating sources, generation requirements, interruptible loads, load curtailment, voltage reduction, and emergency backup through tie lines.

Generation Commitment and Control. This consists of a set of programs used to optimize the production and delivery of power for fuel savings.

Automatic generation control regulates generator output to match the load in order to maintain the frequency in an area and the sum of all active tie-line power exchanges with neighboring power systems. This program executes cyclically every 2 s.

Economic dispatch minimizes the cost of meeting the energy requirements of the system over a period of time and in a manner consistent with reliable service. Desired generator settings are computed and fed to the automatic generation control program. This cyclic program executes every 5 min and also on demand.

Load forecasting computes the total system hourly load for a specified number of days. It provides an adaptive forecasting system based on observed values of demand and estimated weather conditions. The program generally consists of three mathematical models. A load forecasting model uses past load data to compute hourly load forecasts. A weather forecasting model computes hourly weather forecasts based on past weather history. A weather correction model uses telemetered values of load and weather conditions to correct the forecast. The dispatcher may optionally enter these values through a display.

The historical load and weather data are stored and maintained in files which also contain special events that affect load such as holidays or unusual weather phenomena.

Unit commitment determines a schedule for optimal startup and shutdown of thermal units which minimizes unit startup costs subject to generation objectives, predicted area requirements (load forecast), security (spinning reserve requirements and off-system capacity), and operational constraints (unit minimum up and down times, limits, ramp rates, maintenance and derating schedule). It minimizes the operating costs of the dispatchable generating units. The total dispatchable generation is the sum of the load forecast and net scheduled interchange minus the total nondispatchable generation. The operating cost is defined as the sum of production, startup, shutdown, and maintenance costs. Production cost is calculated by the use of input and output curves adjusted by fuel prices. Transmission losses are also considered using one or more sets of penalty factors. Transmission limits are considered with simple approximations.

The output of unit commitment is the hourly unit schedule. Several sets of schedules (or strategies) may be output for operator review and selection. The output is stored and made available to other applications such as transaction evaluation. This program is executed two times a day.

Reserve monitor calculates the available operating reserve necessary to meet company operating policies. Typically, a spinning reserve, 30-min reserve, and 2-h reserve are computed. This program executes cyclically every 4 min or on demand. The dispatcher is alerted when there is inadequate reserve.

Maintenance scheduling assists operating personnel in scheduling generator maintenance in an optimum manner. It generates a maintenance schedule while taking into account constraints and limitations on available resources required to perform maintenance work. The maintenance schedule indicates when particular generating units will be out of service (unit outages). Maintenance is performed routinely at desired intervals or may be required because of unexpected forced outages.

During the maintenance of a given generator, the outage of this unit is compensated for by other capacities in the power system. This is done by maintaining a level of system reserves. The scheduling of generator outages is performed in a manner which maintains a flat megawatt reserve level or a consistent level of probabilistic level of risk, i.e., probability of emergency procedures.

Interchange Management. Formal and informal interchange agreements may exist between neighboring companies. Operating as an interconnected system has economic and security advantages. While many power pools have control centers to manage the interconnection, most companies prefer to manage, or at least monitor, their own interchange of power. The interchange management programs provide the dispatcher with the ability to make good deals with the neighbors. These interactive routines allow the dispatcher to evaluate a buy or sell transaction before being committed to it. Once initiated, the interchange scheduling program automatically schedules the transaction through the automatic generation control program.

Production costing provides the capability to compare a change in interchange with the current schedule. The program always starts by accessing a study file. The calculations are based on current or forecast system load and a proposed interchange schedule entered by the operator.

Transaction evaluation calculates the costs and savings associated with the sale and purchase of power with a selected interconnected company. The program can be used to evaluate past, present, and future operation costs. A production cost calculation is generally included as a part of this program or may be an independent routine.

For evaluating transactions, the transaction evaluation program computes the costs of savings of a proposed transaction by comparing the production costs computed from two economic dispatch calculations: a base economic dispatch calculation for the operating conditions with and without the proposed transaction.

The transaction evaluation program usually can be executed in two modes which specify the starting point for all calculations. These two modes are whether the unit commitment program will be called or not. Essentially, the two options determine how the generation schedule is to be determined. Mode A (also called *economy A*) is executed without unit commitment. Current system load, interchange schedules in study file, and existing unit commitment form the starting conditions. Mode B (also called *economy B*) is executed with unit commitment and is used for transactions in the future.

Interchange scheduling is used to process and display each scheduled transaction with interconnected utilities. It computes the total net scheduled interchange as a function of time for use by the automatic generation control program. This calculation considers the start and stop times, generator ramp rate, magnitude, and direction of each active interchange transaction. Interchange schedules and cost information may be logged and displayed.

Interchange accounting provides the capability to account for electric power in the system. This electric power, in the form of measured, calculated, or scheduled megawatts, includes megawatts which are generated, consumed, lost, passed through, sold, and purchased. It includes functions for logging, displaying, reporting, and updating data recorded by the realtime system.

System Security. These programs help to reduce the chance of a major outage or blackout condition. They operate on a study database which has been cleansed of missing data and metering errors by the network status and state estimation programs.

Network status determines the current configuration of the network based on the status of circuit breakers and disconnect switches obtained from the real-time data and from manually entered status information for devices not telemetered. The program is executed periodically, whenever a status change is detected, or on dispatcher request. In addition, the program develops the corresponding mathematical model of the network using the impedance data for the current base load flow or state estimation case. The model will be used in the subsequent calculation of real-time system conditions using the load flow or state estimation programs.

State estimation determines the current state of the power system, including voltage levels and power flows, and calculates loss factors for use by the economic dispatch program for generation scheduling. It filters real-time measurements to detect and eliminate known errors; estimates expected values of the next real-time measurements; and determines the current network configuration. It corrects the mathematical model of the current network configuration. Thus the state estimation program acts as a filter between the raw real-time data and the real-time data requirements of other security applications. The state estimation program is executed whenever the network status program detects a change, periodically, or on dispatcher request.

Contingency analysis computes the potential effect of contingencies involving the loss of generation and transmission facilities. A specific set of predefined contingencies is analyzed on a cyclic basis. It simulates a contingency and calculates the changes in bus voltages and power flows resulting from the contingency. The base conditions for this calculation are the bus voltages or power flows obtained from the load flow program.

Interactive load flow (ILF) allows the dispatcher to perform load flow studies for a scheduled outage or analyze corrective actions after an unexpected outage. ILF uses real-time data to project bus loads in a network. A bus is a connection where power lines change direction or voltage. A bus load is the load at this point. The ILF program stores mathematical models of the present and planned networks. A color display terminal is used to display one-line diagrams of these models. Colors are used to differentiate voltages, heavily loaded lines, bus voltages outside of limits, and open circuit lines. The dispatcher uses a light pen or mouse together with the terminal keyboard to retrieve previous load flow cases, make modifications, create displays, execute load flow cases, and store them.

Training simulator uses a model of the power network to produce realistic reactions to a dispatcher in training. This function is usually run on the backup central processing unit (CPU). The program can operate in either of two modes: monitor or simulation mode. In monitor or playback mode, the simulator reflects the changing status of the power system during the time period when a playback tape was written. The trainee can view displays, thus monitoring power system changes, but cannot take any control action.

In simulator mode, the trainer can input system changes, and the trainee is allowed to perform control actions. The simulator software will reflect the state of the power system based upon the changes input and control actions taken. In this manner the trainee can operate the control system without affecting the true state of the power system.

Customer response systems control centers are a critical link in responding to customer complaints and reacting to outages. Customer information systems (CIS), used for many business and commercial purposes, are increasingly being combined with the engineering and operations information in order to speed repair crew dispatch and reduce outage times. Work orders, repairs, and inspections are increasingly being merged with inventory, warehouse, and engineering records to reduce costs, speed reaction times, and improve technical performance.

25.3.3 Power Plant Monitoring and Control

This application approaches the traditional process control application. Several applications such as fuel monitoring and performance calculations increase the requirement for the data processing resource. The three major computer systems for fossil power plant monitoring and control are described below.

Process Control System. This is a closed-loop control system which takes its direction from EMS and automatically collects plant data by reading instruments. Physical and electrical parameters associated with the boiler, turbine, and generator are monitored on a continuous cyclic basis. Alarms and events are logged, and control of pumps, valves, and switches for routine functions and for startup and shutdown are provided.

Plant Monitoring System. This is strictly a data-collection system for fuel monitoring, performance calculations, and balance-of-plant calculations; no control actions are performed. Data are stored and retrieved as required to prepare reports and perform analysis. These reports include those required by the plant management, load dispatchers, and planning and engineering groups. Periodic reports are prepared to reflect plant operation. Unit incremental generation cost is determined periodically by collecting data that continuously reflect actual operating conditions. This information is transmitted to the dispatcher for use in load dispatching.

Operational Monitoring System. This is used by plant operators to enter manually collected operational data for record keeping, report writing, and engineering analysis. In addition to these systems, the power plant also may use mini- or microcomputers for security systems, environmental systems, controlled access systems, and chemical analysis systems.

25.3.4 Power Plant Maintenance

Power plant maintenance stores pertinent plant maintenance information for analysis of maintenance costs and evaluation of equipment performance. The interactive portion of the system provides power plant personnel with the capability to enter problem data, planning data, and work execution data. Approval and verification functions are at each step of a work order's progress from problem description through work completion and commitment to history. Interactive functions also are provided for entry and maintenance of an equipment database and for access to equipment history.

The batch portion of the system provides for moving completed and rejected work from the active work database to the equipment history database. Batch functions are also used for work backlogs, scheduled work, and other reports.

Because of its varied data requirements, the maintenance information system also has interfaces to other computer systems in the power company. These are the materials-information system for equipment stocking levels, the personnel information system for labor resources, and the general accounting system for cost tracking. Additionally, text processing services are frequently used.

25.3.5 Fuel Management

Fuel is the single largest expenditure in power-plant operations. Operating support personnel must plan for both short-term (1 year) and long-term (5+ years) fuel availability. Thus control and improvement of fuel cost represent the most substantial contribution to the overall economy of power production. The system is used to administer the procurement of coal and oil. It facilitates the monitoring, reconciliation, and performance of fuel contracts. Accounting functions are used for fuel purchases, transportation costs, fuel usage, and inventory value. Also included are short- and long-term cash flow projections and managerial and regulatory reporting capabilities. The overall fuel management application is an integrated model of the load forecast, generation scheduling and dispatch, and fuel allocation with an objective of optimizing fuel contracts. This application is used routinely with growth and on demand as system perturbations occur.

25.3.6 Load Management

The objective of the load management application is to improve the load factor (ratio of average load to peak load) and to be able to shed selected load during emergency conditions. Most utilities have pilot tested some form of load management. The pilot solutions range from consumer guidance (voluntary) to complete control and metering of the consumer's load.

The benefits of load management have been shown to be enhanced by combining two-way utility to customer communications and control into a distribution automation system. An automated distribution system can have valuable additional functions such as automated meter reading, performance monitoring, and quality-of-service improvement.

Load Curtailment System. This scheme employs one master computer at the power company and several computers at major industrial and commercial load centers. In this scheme there is a limited sensor-based data-acquisition requirement and no closed-loop control. Reduction of load will be manual.

Automatic Meter Reading and Load Management. This hierarchical approach employs the corporate customer information system and the EMS system to direct and receive the meter readings, an intermediate level to serve as a communication concentrator, and an intelligent data control unit to scan remote transponders at the meter locations.

25.3.7 Nuclear Data Center

Functionally, the nuclear data center consists of three systems (Table 25-2), which are described below.

Operational System. This consists of real-time and on-line monitoring of various subsystems in a nuclear plant. The plant process computer system in a nuclear plant is usually provided by the reactor vendor and integrated with the reactor. Additional monitoring facilities are now required by the Nuclear Regulatory Commission (NRC). One specific monitoring system is named the *emergency response information system*, which is now a firm requirement for all operating nuclear plants and a license requirement for plants under construction. It has the following components:

- *Safety parameter display system*. This provides additional monitoring and display facilities to the primary plant control room. The additional facilities provide continuous monitoring of critical parameters.
- *Technical support center*. This is a separate facility and staff within the plant to assist the primary plant control room during emergency conditions.

TABLE 25-2 Nuclear Data Center Functions

Operational	
Plant process computer system	
Operator training simulator	
Plant security system	
Emergency response information system	
Radiation monitoring system	
Technical	
Fuel shuffling	Radioactive waste control
Fuel management	Meteorological/environment
Refueling outage scheduling	Chemical laboratory
Quality assurance	Health physics
Start-up testing	Technical services
Administrative	
Maintenance	Nuclear records management
Materials management	Financial control
Licensing	Plant operation
Document control	Personnel
	Training

- *Emergency operations facility*. This is a separate facility and staff located off the plant site but in proximity to the plant. It will manage the overall emergency response and evaluate actual and potential radiological releases.
- *Nuclear data link*. This routes cyclic data from the plant to the NRC headquarters through a communication link. The critical data will be used to monitor an accident, to evaluate emergency conditions, to advise and assist the licensee, and to inform the general public.

Administrative System. This is used for maintenance and backup of the operational system. Storage and information retrieval systems are used for the storage and retrieval of large amounts of data as encountered in nuclear records management. These systems provide interactive direct access to, and fast scanning or searching of, large numbers of articles, reports, contracts, laws, general directives, or abstracts of publications.

Technical System. This is used for large simulation and nuclear fuel management programs. Computer programs simulating nuclear core physics in various levels of detail are presently being used for analysis, leading to continual improvement and enhancement of nuclear plants. These programs are used not only in the design of cores for nuclear reactors but also for the planning and decision making regarding fuel burnup and associated electrical generation for nuclear units. Core refueling schedules are of vital concern. Safety analysis for emergency situations is another area in which the simulation programs are used. Nuclear fuel management programs are used for the design of nuclear reactors and for nuclear fuel management calculations.

25.4 ENGINEERING COMPUTING TRENDS

Costs. Today the cost of the engineering user is 8 to 50 times the cost of the computer that the engineer uses. This range depends on the level of the user and the type of processing: machine-intensive (compiling, executing) or human-intensive (editing, scrolling). It is no longer cost-effective to economize on computing resources required by the engineering user.

Personal Computing. Batch-job and central processing is rapidly being replaced by networked or individual personal computers and workstations. The fast performance and low cost of this equipment are a perfect complement to modern interactive, graphics-oriented software, which combine to greatly improve engineering productivity and effectiveness.

Networks. High-speed networks allow easy communication of data, results, and messages, leading to an integrated set of applications spanning functions and levels of the company. Ethernet and token-ring systems and communication protocols such as TCP/IP provide versatile and very high speed communications between diverse computers.

Languages. Most engineering computer applications are written in FORTRAN because of the long-time popularity of this language. New applications and graphics user interfaces (GUI) are usually written in C or the object-oriented C++ language. Graphics interfaces use special facilities such as Windows, MOTIF, and OpenLook.

On-Line Reporting. This has replaced manual reporting of data. Although text processing systems are used to generate documents, text information can also be transmitted, distributed, and stored electronically to reduce the proliferation of paper documents.

Results Selection and Editing. This is now performed by scrolling a display instead of printing all results and sifting through reams of paper.

Automated Data Flow. The trend is to maintain one set of data that is available to all users. These data are protected by appropriate security features and procedures like any other corporate resource. However, the intent is to make common data available to all users and allow the appropriate users to update the common data. This common database leads to accuracy, avoids duplication, and avoids loss of labor due to the reentry of data.

Productivity Aids. While most of the computer applications that the engineer uses are complex programs, the productivity of the engineer is also increased by simple, direct tools for text processing, graphics, and project control. These tools are interactive and require a minimum of effort to use. These tools help the engineer make notes, drafts, lists, sketches, and establish personal meeting or project schedules.

Computer Configurations. Computer configurations such as shown in Figs. 25-8 and 25-9 are used to support engineering functions. Although the centralized computing approach has traditionally been used, the trend is to a distributed approach in which a separate engineering information center is assigned strictly to engineering work.

Engineering User Service Objectives. Engineering users require three categories of computing support:

1. *Performance requirements.* Less than 0.4-s response time is required for human-intensive work at a terminal. Longer response times reduce concentration and increase frustration. A predictable response time is required for machine-intensive work. Consistent timings for load flow are expected when operating under deadlines. High-bandwidth terminals are required to support Windows graphics user interface (GUI) color graphics applications, which are used to speed work flow and to reduce volumes of data to a presentation of results in a form that can be readily assimilated.

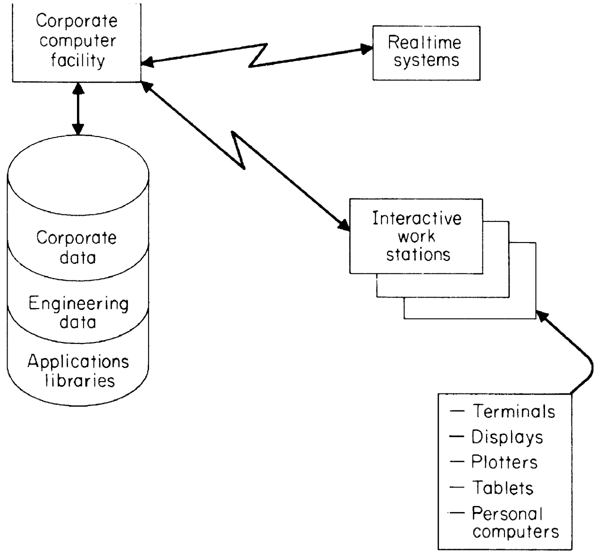


FIGURE 25-8 Power engineering centralized architecture.

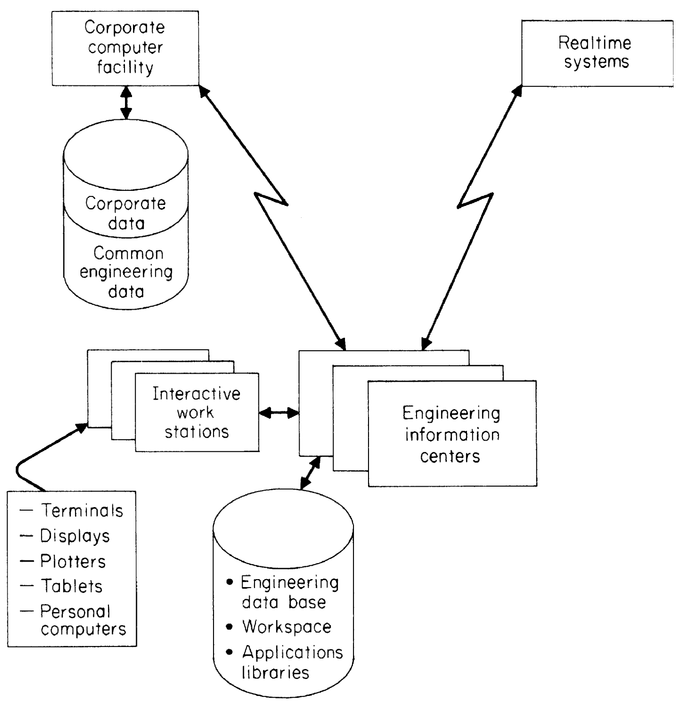


FIGURE 25-9 Power engineering distributed architecture.

2. *Technical support requirements*. The engineering user requires a minimum of system programming knowledge to use the computing system. Included in this are user-friendly selection panels, menus, aids, and help functions; a library of programming tools, languages, and application packages, all of which are maintained by the central data processing department; and training facilities and technical support to assist engineers with the use of computing resources.
3. *Operational support requirements*. Security features and operational procedures are required to protect the user's data and access to the computing system. Included are security features such as password protection of user data files and sign-on's and operational procedures to provide backup of user data and prevent its destruction or loss.

BIBLIOGRAPHY

Ackerman, W. J., Evans, J. W., and Smith, H. L., *Adding New Life To Legacy SCADA Systems*, IEEE Tutorial (03TP163), Power Engineering Society (PES) General Meeting, Toronto, 2003.

Maghsoodlou, F., Masiello, R., and Ray, T., *Energy Management System*, IEEE Power & Energy Magazine, September/October 2004.

EPRI, *Power Market Simulator for Wholesale Energy Markets*, Palo Alto, California, 2005, www.epri.com

SECTION 26
ILLUMINATION*

Richard G. Mistrick
Associate Professor, Penn State University; Professional Engineer; Lighting Consultant; Fellow, Illuminating Engineering Society of North America (IESNA); MEMBER, U.S. National Committee of the International Commission on Illumination.

CONTENTS

26.1 RADIANT ENERGY AND LIGHT

For the principal purposes of illumination design, light is defined as visually evaluated radiant energy. The visible energy radiated by light source is found in a narrow band in the electromagnetic spectrum (Fig 26-1) approximately from 380 to 770 nanometers (nm). By extension, the art and science of illumination also include the applications of ultraviolet and infrared radiation. The principles of measurement, methods of control, and fundamentals of lighting system and equipment design in these fields are closely parallel to those long established in lighting practice.

26.2 QUANTITIES, UNITS, AND CONVERSION FACTORS[1]

Luminous Flux. This is the time rate of flow of light. See Table 26-1. Radiant energy in the visible region of the spectrum varies in its ability to produce visual sensation, the variation depending upon

*Includes some material from previous editions by Jack F. Parsons, Walter Sturrock, Karl A. Staley, John A. Kaufman, and Charles Amick.

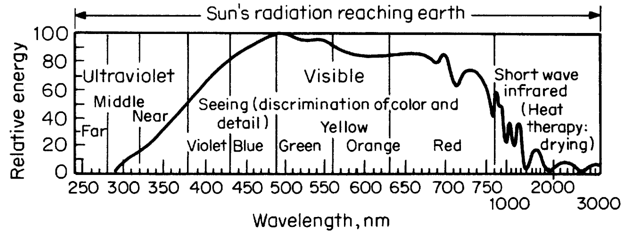


FIGURE 26-1 Ultraviolet, visible, and shortwave infrared are the three principal bands of the electromagnetic spectrum with which illuminating engineering is concerned.

the wavelength. The ratio of the luminous flux to the corresponding radiant flux is known as spectral luminous efficacy and is expressed in lumens per watt (lm/W). This varies with wavelength, having a maximum at approximately 555 nm. The data are plotted in Fig. 26-2. At very low levels of illumination the position of the maximum sensitivity gradually shifts to 510 nm as a result of greater use of rod vision.

From the foregoing it is apparent that two sources may radiate equal amounts of energy in the visible region of the spectrum but have different amounts of luminous flux emitted, depending on the spectral distribution of the energy. The luminous flux (ϕ) is the integrated product of the energy per unit wavelength emitted by the source $P(\lambda)$, referred to as the source's spectral power distribution, and the spectral luminous efficacy $V(\lambda)$ as follows:

$$\phi = 683\int_{\lambda=360}^{800} P(\lambda)V(\lambda)d\lambda$$

The lumen is the unit of luminous flux. Light sources (i.e., lamps) are rated in lumens.

Luminous Intensity. This is the luminous flux per unit solid angle in a specific direction. Hence, it is the luminous flux on a small surface normal to that direction, divided by the solid angle (in steradians) that the surface subtends at the source (see Table 26-1). The definition of luminous

TABLE 26-1 Standard Units, Symbols; and Defining Equations for Fundamental Photometric Quantities

Quantity*	Symbol	Defining equation	Unit	Symbolic abbreviation
Luminous flux	ϕ	$\Phi = dQ/dt$	lumen	lm
Illuminance (illumination)	E	$E = d\phi/dA$	footcandle (lumen per Φ square foot)	fc
			lux (lm/m^2)	lx
Luminous existance	M	$M = d\Phi/dA$	lumen per square foot	lm/ft^2
Luminous intensity (candlepower)	I	$I = d\Phi/d\omega$ (ω = solid angle through which flux from point source is radiated)	candela (lumen per steradian)	cd
Luminance (photometric brightness)	L	$L = d^2\Phi/d(dA\ cos\ \theta)$ $= dI/(dA\ \cos\theta)$ (θ = angle between line of sight and normal to surface considered)	candela per unit area	ω etc.
			cd/ft^2 nit (ed/m^2)	cd/ft^2
			footlambert ($cd/\pi ft^2$)†	fL
Luminous efficacy	K	$K = \Phi_r/\Phi_\rho$	lumen per watt	lm/W

*Quantities may be restricted to a narrow wavelength band by adding the word s*pectral* and indicating the wavelength. The corresponding symbols are changed by adding a subscript λ, e.g., Q_λ, for a spectral concentration or a λ in parentheses, e.g., $K(\lambda)$, for a function of wavelength.

†The use of this unit is deprecated.

intensity applies strictly to a point light source. In practice, however, light emanating from a source whose dimensions are negligible in comparison with the distance from which it is observed may be considered as coming from a point.

Candlepower is another term for luminous intensity, since the candela is the unit of luminous intensity. One candela is defined as the luminous intensity of 1/600,000 $m^2$of projected area of blackbody radiator operating at the temperature of solidification of platinum under a pressure of 101,325 Pa. It is also the luminous intensity when one lumen is directed within one steradian of solid angle. A steradian is a unit area on a sphere of radius one, thus there are 12.57 (4π) steradians surrounding any light source. The original definition of luminous intensity was in terms of the strength of a flame source, a standard candle.

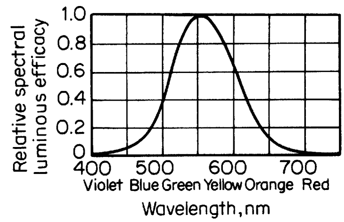


FIGURE 26-2 Spectral luminous efficacy, $V(\lambda)$, for normal human color vision.

Illuminance. This is the density of the luminous flux incident on a surface; it is the quotient of the luminous flux by the area of the surface when the latter is uniformly illuminated. The term illumination is used to designate the act of illuminating or the state of being illuminated. Usually the context will indicate which meaning is intended, but the expression "level of illumination" is a term used to mean illuminance and should be discouraged.

Lux is the unit of illuminance when the meter is taken as the unit of length. It is the illumination on a surface 1 m^2 in area on which there is a uniformly distributed flux of 1 lm, or the illumination produced on a surface, all points of which are at a distance of 1 m form a directionally uniform point source of one candela.

Footcandle is the inch-pound system unit of illuminance where the foot is taken as the unit of length. See Table 26-2 for conversion factors between SI and inch-pound lighting units. Most conversions, like illuminance, involve a 10.76 factor since there are 10.76 ft^2/m^2 (e.g., 1 footcandle = 10.76 lux).

Luminous Exitance. This is the density of luminous flux leaving a surface; it is the quotient of the luminous flux leaving the surface by the area of the surface. It applies to the aggregate flux that is emitted, reflected, or transmitted from the surface and is a nondirectional quantity.

Luminance. This is the quotient of the luminous flux leaving or arriving at an element of a surface and propagated in direction defined by an elementary cone containing the given direction, by the

TABLE 26-2 Conversion Factors for Lighting Units

A. Illuminance			
1 footcandle = 1 lumen per square foot		1 lux = 1 lumen per square meter = 1 meter-candela	
		Footcandles	Lux
Footcandles		1	0.0929
Lux		10.76	1
B. Luminance			
1 footlambert = 1 lumen per square foot		1 nit = 1 candela per square meter	
	Foot-lamberts	Candelas per square meter	Candelas per square foot
Footlamberts	1	0.2919	3,142
Candelas per sq m	3,426	1	10.76
Candelas per sq ft	0.3183	0.0929	1

product of the solid angle of the cone, and the area of the orthogonal projection of the element of the surface on a plane perpendicular to the given direction. More simply, it is the luminous intensity of any surface in a given direction per unit of projected area of the surface viewed from that direction (see Table 26-1).

Candela per square meter is the SI unit of luminance when the meter is taken as the unit of length. Another term for this unit is the nit, which is not commonly used in North America. The candela per square foot is the inch-pound unit of luminance when the foot is the unit of length.

Footlambert is a former unit of luminance, and is equal to $1/\pi cd/ft^2$, or to the uniform luminance of a perfectly diffusing surface emitting or reflecting light at the rate of 1 lm/ft^2, or to the average luminance of any surface emitting or reflecting light at that rate. The term footlambert is now obsolete, and its use is deprecated.

Luminous Efficacy. This is a quantity denoting the energy effectiveness of light sources. It is the ratio of the total luminous flux (lumens) to the total power input (watts). The maximum luminous efficacy of an *ideal* white source, defined as a radiator with constant output over the visible spectrum, is approximately 200 lm/W.

Reflectance. Reflectance ρ is the ratio of reflected flux to incident flux. Measured values of reflectance depend upon the angles of incidence and view, and on the spectral character of the incident flux. Because of the dependence, the angles of incidence and view and this spectral characteristics of the source should be specified.

Transmittance. Transmittance τ is the ratio of the transmitted flux to the incident flux. Measured values of transmittance depend upon the angle of incidence, the method of measurement of the transmitted flux, and the spectral character of the incident flux. Because of this dependence, complete information of the technique and conditions of measurement should be specified.

Absorptance. Absorptance α is the ratio of the flux absorbed by a medium to the incident flux. The sum of reflectance, transmittance, and absorptance is one.

Brightness. This term refers to the intensity of sensation resulting from viewing light source and surfaces. This sensation is determined in part by the measurable luminance defined above and in part by conditions of observation such as the state of adaptation of the eye.

Color. Within the visible spectrum, wavelengths are distinguished one from another by their ability to excite in the human eye various color sensations. Thus the shorter wavelengths excite the color known as violet, and as the wavelengths increase, the color sensation gradually changes through blue, green, yellow, and orange, and finally to red at the longer wavelengths of the visible spectrum. The color of the sensation produced by light of a composite character is determined by its spectral power distribution. Color is defined as that quality of visual sensation which is associated with the spectral distribution of light. Color matching is the process of adjusting the color of one area so that it is the same color as another.

Correlated color temperature (CCT) of a light source is the absolute temperature (in Kelvin) of a blackbody radiator whose chromaticity most nearly resembles that of the light source. CCT refers to the whiteness of the light that a source emits. Low CCT light appears more yellow or red, and is generally considered to be warm in appearance, while high CCT light appears more bluish white. A neutral CCT is generally considered to be around 3500 K.

Color rendering is a general expression for the effect of a light source on the color appearance of objects in conscious or subconscious comparison with their color appearance under a reference light source.

The Color rendering index (CRI) of a light source is the measure of the degree of color shift which objects undergo when illuminated by the light source, as compared with the color of those same objects when illuminated by a reference source of comparable color temperature. Values for common light source vary from about 20 to 99. The higher the number, the better the color rendering

TABLE 26-3 Color Temperature and Color Rendition Index of Some Common Light Sources*

Light source	Correlated color temperature, K	Color rendering index
"Cool" fluorescent		
Standard cool white ES†	4150	62
Cool white, ES, RE741, phosphor	4100	72
Lite white, ES	4200	49
RE841 phosphor	4100	80
"Warm" fluorescent		
Standard warm white, ES	3000	52
Warm white, ES, RE730 phosphor	3000	70
RE830 phosphor	3000	82
RE827	2700	82
Deluxe daylight fluorescent, ES	6500	84
RE950	5000	90
Incandescent		
General service	2600–3100	89–92
Tungsten-halogen	2900–3100	90
High-intensity discharge		
Mercury	5710	15
Mercury improved color	4430	32
Metal halide, clear	4000	65
Metal halide, ceramic	3000	80-88
Metal halide, ceramic	4200	90
High-pressure sodium	2100	21
Daylight		
Overcast sky	6000–7000	
Blue sky	11,000–25,000	
Sun, outside of earth's atmosphere	6500	

*Check manufacturer's technical literature for current data.
†Energy-saving models.

(see Table 26-3). CRI should only be used to compare sources of the same color temperature since different reference sources are used at different color temperature. A black body radiator is used at low CCT's and daylight spectra are used at high CCT's.

REFERENCE ON QUANTITIES, UNITS, AND CONVERSION FACTORS

1. *American National Standard Nomenclature and Definitions for Illuminating Engineering, RP-16-05, Illuminating Engineering Society of North America.*

26.3 INCANDESCENT LAMPS

Incandescent Filament Lamps. These are light sources in which light is produced by a filament heated to incandescence by an electric current. Of all commonly used light sources, incandescent lamps have the lowest initial cost, lowest luminous efficacy, and shortest life. As shown in Fig. 26-3, the major parts on an incandescent filament lamp are the filament, bulb, base and fill gas.

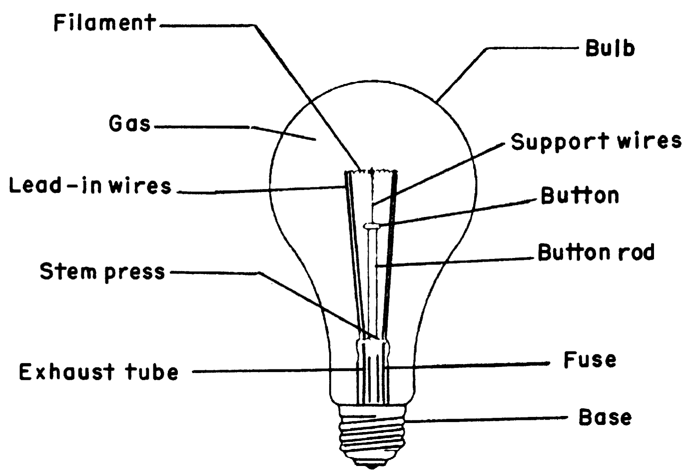


FIGURE 26-3 Incandescent filament lamp construction.

Filament. The efficacy of light production by incandescent lamps depends on the temperature of the filament. Tungsten, because of its high melting point (3655 K), higher than that of all other elements except carbon, is the most common filament material used today. Filament forms, sizes, and support constructions vary with different types of lamps. Most filaments are coiled one or more times to increase the filament temperature, light output, and the lamp's luminous efficacy.

Mechanical problems associated with tungsten filaments make the incandescent lamp an inherently compact, somewhat spherical structure. The filament's length and diameter limit its range of operation between 1.5 and 300 V. At 1.5 V, the filament is very short and thick, and it becomes difficult to heat it without excessively heating its support wires. The lamps in the low-voltage (6-to 12-V) class, however, are relatively rugged and will withstand the shocks of motor-vehicle and similar applications. At voltages near 30 V, the filament is very long and slender; it is fragile and difficult to support.

Bulbs. Bulb shape, size, material, and finish vary finish vary according to application needs. Shapers range from tubular to spherical and from parabolic to flame form. Bulbs are designated by a letter referring to the shape (see Fig.26-4) and by a number which is the maximum diameter in eights of an inch; for example, A-19 designates an A-shaped bulb with a diameter of 19/8 or 2-3/8 in.

Most bulbs are made of lead or lime soft glass, although heat-resisting hard glass is used for high-temperature application, and are frosted on the inside for moderate diffusion of the light without appreciably reducing light output. Clear, unfrosted lamps are used where accurate control of light is needed from a point or line source. Fused quartz and high-silica glass are used for other lamps.

Base Types. These also very according to application needs. They range from screw types for most general-service lamps to bipost and prefocus types where a high degree of accuracy in lamp positioning is important, such as in projection systems. Figure 26-5 shows some typical base shapes. Base size varies with lamp wattage, for heat dissipation, and voltage. For outdoor lighting, use brass-base lamps.

Fill Gas. This is used in incandescent-filament lamps to reduce the rate of evaporation of the heated filament. Inert gases such as nitrogen, argon, and krypton are in common use today, with krypton used where its increased cost is justified by increased efficacy or increased lamp life. For example, the 90-W krypton "energy-saving" lamp produces 4% less light, but one-third longer rated life compared with the standard 100-W lamp.

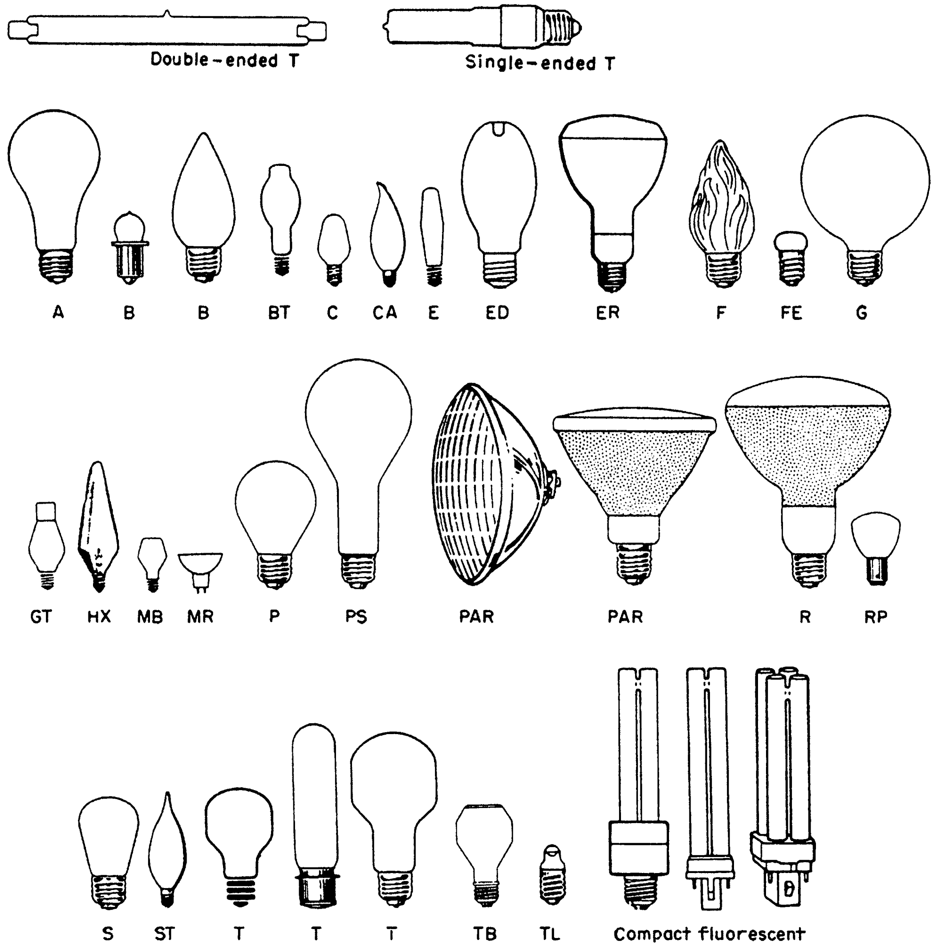


FIGURE 26-4 Typical bulb shapes and designations (not to scale). Most high intensity discharge (HID) lamps use BT-, E-, ED-, PAR-, and R-shape bulbs.

Many regular, tubular, and PAR shaped lamps are available with a halogen fill gas, for better lumen maintenance, improved light output, and/or longer life. Called "tungsten halogen," their filaments operate at temperatures higher than regular incandescent lamps, producing light of greater color temperature, plus longer life for a given light output. For applications justifying the increased lamp cost, a 90-W, 2000-h tungsten halogen lamp, for example, has only 8% less light output rating than the standard 100-W, 750-h lamp, with improved lumen maintenance through its longer life.

Energy Characteristics. Only a small percentage of the total radiation from incandescent lamps is in the visible spectrum, with the majority in the infrared spectrum. As the filament temperature is increased, the luminous efficacy increases with a maximum of 53 lm/W for an uncoiled tungsten wire at its melting point. To obtain life, practical lamps operate at a temperature will below the melting point.

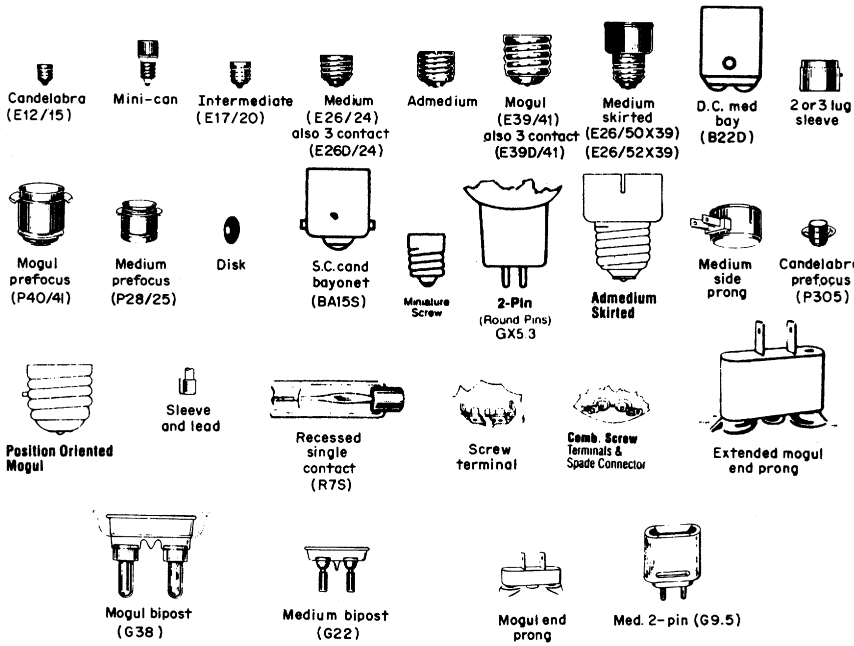


FIGURE 26-5 Common incandescent and HID lamp bases (not to scale). IEC designations are shown where available.

Performance Characteristics. The performance of tungsten-filament lamps is affected by voltage, position of the bulb (if incorrect), size, construction, ambient temperature (if excessive), and quality of manufacture. The voltage characteristics through a range of a few volts above and below design volts may be expressed as simple exponential equations in the following relationships, where capitalized terms represent normal rated values:

$$\frac{\text{Life}}{\text{LIFE}} = \left(\frac{\text{LUMENS}}{\text{lumens}}\right)^a = \left(\frac{\text{LUMENS / WATT}}{\text{lumens / watt}}\right)^b = \left(\frac{\text{VOLTS}}{\text{volts}}\right)^d = \left(\frac{\text{AMPS}}{\text{amps}}\right)^u$$

$$\frac{\text{Lumens}}{\text{LUMENS}} = \left(\frac{\text{volts}}{\text{VOLTS}}\right)^k = \left(\frac{\text{lumens / watt}}{\text{LUMENS / WATT}}\right)^h = \left(\frac{\text{watts}}{\text{WATTS}}\right)^s = \left(\frac{\text{amps}}{\text{AMPS}}\right)^y = \left(\frac{\text{ohms}}{\text{OHMS}}\right)^z$$

$$\frac{\text{LUMENS / WATT}}{\text{lumens / watt}} = \left(\frac{\text{LUMENS}}{\text{lumens}}\right)^f = \left(\frac{\text{VOLTS}}{\text{volts}}\right)^g = \left(\frac{\text{AMPS}}{\text{amps}}\right)^j$$

$$\frac{\text{amps}}{\text{AMPS}} = \left(\frac{\text{volts}}{\text{VOLTS}}\right)^t \quad \text{and} \quad \left(\frac{\text{watts}}{\text{WATTS}}\right) = \left(\frac{\text{volts}}{\text{VOLTS}}\right)^n$$

Exponents *d, k,* and *t* are taken as fundamentals, and other exponents are derived from them. A list of exponents is given in the following table. Values given apply to lamps operated at 90% to 110% rated voltage. Outside that range, use the values from Fig. 26-6.

The theoretical life of lamps calculated by the exponential relationship of life and voltage is seldom realized in practical installations in the case of excessive "undervoltage" burning, since handling, cleaning, vibration, etc., introduce breakage factors which tend to reduce lamp life.

Exponents

	Gas-filled	Vacuum
a	3.86	3.85
b	7.1	7.0
d	13.1	13.5
u	24.1	23.3
k	3.38	3.51
h	1.84	1.82
s	2.19	2.22
y	6.25	6.05
z	7.36	8.36
f	0.544	0.550
g	1.84	1.93
j	3.40	3.33
t	0.541	0.580
n	1.54	1.58

Lamp Lumen Depreciation. Because of filament evaporation throughout life, the filament of a lamp becomes thinner and thus consumes less power. The light output decreases as the lamp progresses through life because of lower filament temperature and bulb blackening. Figure 27-7a shows the change in watts, amperes, lumens per watt, and lumens for a 200-W general-service lamp on constant-voltage service. The minor quantity of bromine or iodine in tungsten-halogen lamps vaporizes during operation, and acts to return particles of tungsten back to the filament. This results in superior lumen maintenance.

Lamp Mortality and Renewal Rate. Lamp life is based on data obtained form lifetesting a large number of lamps. A perfect mortality record would be one in which all lamps reached their rated life and then burned out. However, many factors inherent in lamp manufacture and lamp materials make it impossible for each individual lamp to operate for exactly the life for which it was designed. A typical mortality curve of a large group of lamps is illustrated in Fig. 26-7*b*, where it is superimposed on a lumen depreciation curve from Fig. 26-7*a*.

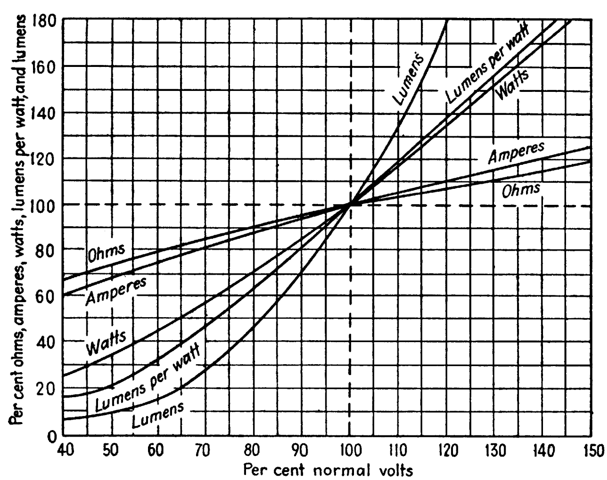


FIGURE 26-6 Characteristic curves for large gas-filled lamps showing the effect of operating a lamp at other than its rated voltage. These characteristic curves are averages of many lamps.

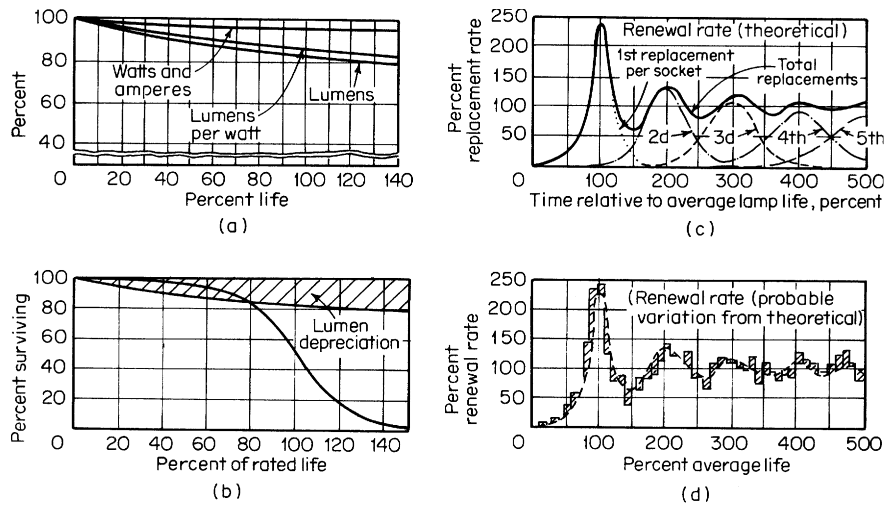


FIGURE 26-7 Life characteristics and renewal rate.

The mortality curve influences the rate of lamp replacements for installations involving a large number of lamps. If individual lamps are replaced as they burn out, the replacement rate is as shown in Fig. 26-7*c*. In a new installation relatively few burnouts would be expected during the first several hundred hours of operation, but as the design life is approached, the rate of burnout increases rapidly. After a burning period of 4 to 5 times the average lamp life, the renewal rate fluctuation finally reaches a steady or normal rate.

The dotted curve in Fig. 26-7*c* showing the theoretical rate of renewals holds only for an infinitely large installation. Departures from this curve in practical installations will, by the law of probability, more likely be represented by the solid block-shaped pattern. The larger the installation, the more closely the two curves tend to coincide. Complaints on life are occasionally encountered during those periods when chance dictates that renewals run higher than average, even though a record of the actual number of renewals over an extended period of time would show average rated life had been obtained.

Rated Lamp Life. The rated life of a lamp is generally defined as the operating hours at which 50% of a representative group of lamp burned under correct operating conditions on a 60-Hz circuit are still burning, and ranges from 750 to 1500 h for the general-service incandescent types. As compared with life in the laboratory under controlled operating conditions, performance in service may differ widely. Lamp breakage and fluctuating line voltage tend to shorten life. Line-potential drop with resultant low-voltage operation often tends to lengthen life. Extended-service lamps with a rated of 2500 h and longer are available in a range of sizes from 15 to 1000 W. They give less light than standard lamps under normal conditions but may be economically justified when labor costs to replace lamps are very high.

Influence of operating Conditions on Lamp Performance. Tests show that ambient temperatures have little effect on performance characteristics. Very high temperatures, however, may cause mechanical difficulties. On direct current, although the mortality rate is lower, the maintenance of light output is poorer than on alternating current. Intermittent operation in general (not sign-flashing service) does not materials affect lamp performance. There is a reason to believe that lamp life is shortened by voltage fluctuations, even though the voltage excess averaged over the life of the lamp is offset by an equal average voltage deficiency.

Except in the case of lamps designed for a particular position of operation, operating position has little effect on lamp performance. Shock and vibration are likely to impair the performance of lamps with filaments of small diameter to a greater extent than in the case of lamps with filaments of large diameter. Special types of lamps are available for use in installations where vibration is likely to be encountered and others, known as "rough-service" lamps, for use where they are likely to be subjected to shock. Neither of these two lamps will function properly in place of the other.

Classes of Incandescent Lamps. Incandescent lamps are divided and cataloged by manufactures into three major groups: large lamps, miniature lamps, and photographic lamps. Large lamps are those normally used for interior and exterior general and task lighting. Miniature lamps are generally used in automotive, aircraft, and appliance applications. Photographic lamps, as the name implies, are used in photography and projection service. Some of the main classes of large lamps are as follows:

General Service. These are for general lighting on 120-V circuits (see Table 26-4). Sizes range from 15 to 1500 W with efficacies of 8 to nearly 23 lm/W. Lamps rated at 130 V and other voltages are also available—consult catalogs or sales representatives of lamp manufactures for specific listings.

High Voltage. High-voltage incandescent lamps are designed for operation directly on circuits of 220 to 300 V. They are less rugged and have a lower efficacy than general-service lamps. There are also general-service incandescent lamps of 277-V circuits. One manufacturer cautions that such lamps be enclosed if used on high-capacity, low impedance electrical distribution systems.

Extended Service. These have a life of 2500 h or more and are intended for use in applications where a lamp failure causes an inconvenience, a nuisance, or a hazard to replace the lamp, or where replacement labor is expensive (see Table 26-5). They are less efficient than general-service lamps.

TABLE 26-4 General-Service Lamps for 120-V Circuits*
(Will operate in any position, but lumen maintenance is best for 40 to 1500 W when burned vertically base up)

Watts	Bulb and other description	Base	Rated average life, h	Maximum overall length, in	Approximate initial lumens	Rated initial lumens per watt	Depreciation factor, % output at 70% rated life
15	A-15 inside frosted	Med.	2500	$3^1/_2$	125	8.3	83
25	A-19 soft white	Med.	2500	$4^1/_2$	210	8.4	
34†	A-19 diffuse coating	Med.	2000	$4^7/_{16}$	380	11.1	
40	A-19 inside frosted	Med.	1000	$4^7/_{16}$	505	12.6	
52†	A-19 diffuse coating or clear	Med.	1330	$4^7/_{16}$	730	14.0	
60	A-19 inside frosted	Med.	1000	$4^7/_{16}$	865	14.4	93
67†	A-19 diffuse coating	Med.	1000	$4^7/_{16}$	1,030	16.9	
75	A-19 inside frosted	Med.	750	$4^7/_{16}$	1,190	15.9	92
90†	A-19 diffuse coating	Med.	1000	$4^7/_{16}$	1,465	16.3	
100	A-19 inside frosted	Med.	750	$4^7/_{16}$	1,710	17.1	90.5
135†	A-21 diffuse coating	Med.	1000	$5^3/_8$	2,380	17.6	
150	A-21 inside frosted or clear	Med.	750	$5^3/_8$	2,850	19.0	89
200	A-23 inside frosted or clear	Med.	750	$6^5/_{16}$	3,920	19.6	89.5
300	PS-25 clear or inside frosted	Med.	750	$6^{15}/_{16}$	6,200	20.7	87.5
300	PS-35 clear or inside frosted	Mog.	1000	$9^3/_8$	5,820	19.4	86
500	PS-35 clear	Mog.	1000	$9^3/_8$	10,850	21.7	89
750	PS-52 clear	Mog.	1000	13	17,040	22.7	89
1000	PS-52 clear or inside frosted	Mog.	1000	13	23,740	23.7	89
1500	PS-52 clear or inside frosted	Mog.	1000	13	34,400	22.9	78

*Consult manufacturers' technical literature for current data, as values change frequently.
†Reduced wattage, krypton-fill type.
Note: 1 in = 25.4 mm.

TABLE 26-5 Extended Service (2500-h Rated Life) Incandescent Filament Lamps*
(For 120 V)

Watts	Bulb and Finish	Base	Maximum overall length, in	Approximate initial lumens	Rated initial lumens per watt	Depreciation factor, % output at 70% rated life
19†	Soft-white	Med.	3 1/2	110	7.3	
25	A-19 soft white	Med.	4 1/4	210	8.4	79
34†	A-19 diffuse coating	Med.	4 7/16	375	11.0	
67†	A-19 diffuse coating	Med.	4 7/16	940	14.0	
90†	A-19 diffuse coating	Med.	4 7/16	1285	14.3	
100	A-19 diffuse coating	Med.	4 7/16	1440	14.4	92.5
135†	A-19 diffuse coating	Med.	5 3/8	2100	15.6	
150	A-23 inside frosted	Med.	6 5/16	2350	15.7	89
150	PS-25 inside frosted	Med.	6 15/16	2250	15.0	85.5
200	A-23 inside frosted and clear	Med.	6 3/16	3320	16.6	87.5
300	PS-30 inside frosted and clear	Med.	8 1/16	5190	17.3	79
300	PS-35 inside frosted and clear	Mog.	9 3/8	5190	17.3	84
500	PS-40 inside frosted and clear	Mog.	9 3/4	9070	18.1	80

*Consult manufacturers' technical literature for current data, as values change frequently.
†Reduced wattage, krypton-fill type. Electrical, light output, and life ratings are different for various manufacturers.
Note: 1 in = 25.4 mm.

General Lighting Tungsten-Halogen. These are compact, have better lumen maintenance, and provide a whiter light and a longer life. Some typical lamps for general lighting are listed in Table 26-6.

Reflectorized. These are a group of lamps embodying integral reflecting surfaces. Bowl-silvered lamps are employed in direct-lighting equipment in which it is desired to shield the filament from view in direct or indirect equipment. Initial loss of light output due to the silvering is 6% to 10%; the rate of decline of light output is considerably greater than in clear-bulb lamps of corresponding sizes—60% to 80% greater in the case of 100- and 200-W lamps. However, a luminaire (of similar distribution) with an unprocessed lamp may produce less light because of poorer maintenance.

In projector flood- and spotlight-lamps, the bulb is constructed of a molded bowl-shaped section of parabolic or other suitable profile, on the inner surface of which is a metal-reflecting surface (see Table 26-6). This bowl is fused to a molded-glass cover plate, which may be clear or may consist of a pattern of lenses and prisms, depending on the desired beam characteristics. Reflector-type lamps are constructed with blown bulbs of suitable profiles (usually cylindrical for showcase lighting or parabolic for spotlighting) having parts of the inner surfaces covered with a reflecting metallic film. Their nominal life is usually 2000 h.

The National Energy Policy Act (EPACT) of 1992 prohibited the manufacture after October 31, 1995 of certain standard, general-service 115- to 130-V reflector and projector lamps. Among those obsoleted were 50-, 75-, and 100-W R-40 lamps and 75-, 100-, and 150-W R-40 and PAR-38 lamps. They are replaced by higher-efficiency halogen-capsule lamps.

A number of projector lamps are available with dichroic filters (interference films) to control the spectral quality of the radiation in such a manner as to separate the heat from the light in the beam or to produce colored light without the usual losses due to absorption by filters. From 75% to 80% of the heat can be removed from the beam at a sacrifice of only 15% to 20% of the light. These "cool beam" lamps must be used in luminaires that are capable of dissipating the additional heat that remains within the luminaire. Colored dichroic lamps produce more deeply saturated colors with higher efficacy than is obtainable with color filters.

In certain PAR-bulb lamps, the tungsten-halogen capsule has an infrared coating, enabling a lower power rating by redirecting energy back to the filament (Table 26-7). All applications of reflectorized lamps should follow the recommendations of the manufacturer concerning luminaire

TABLE 26-6 Tungsten-Halogen Lamps for General Lighting*

Watts	Volts	Bulb and Finish	Base†	Maximum overall length, in	Rated life, h	Approximate initial lumens	Approximate initial lumens per watt	Depreciation factor, % output at 70% rated life
				Double-ended types				
200	120	T-3 clear	RSC	$3\frac{1}{8}$	1500	3,460	17.3	
300	120	T-3 clear	RSC	$4\frac{11}{16}$	2000	5,950	19.9	96
300	120	T-4 clear	RSC	$3\frac{1}{8}$	2000	5,650	18.5	
400	120	T-4 clear	RSC	$3\frac{1}{8}$	2000	7,750	19.4	96
500	120	T-3 clear	RSC	$4\frac{11}{16}$	2000	11,100	22.2	96
1000	120	T-6 clear	RSC	$5\frac{5}{8}$	2000	23,400	23.4	96
1000	220	T-3 clear	RSC	$10\frac{1}{16}$	2000	21,500	21.5	96
1500	220	T-3 clear	RSC	$10\frac{1}{16}$	2000	35,800	23.9	96
				Single-ended types				
50	120	TB-19 I. F.	Med. screw	$4\frac{7}{16}$	2000	710	14.2	
90	120	TB-19 I. F.	Med. screw	$4\frac{7}{16}$	2000	1,580	17.6	
100	120	T-4 clear	Minican	$2\frac{13}{16}$	2000	1,600	16.0	
150	120	T-4 clear	Minican	3	2000	2,800	18.7	
250	120	T-4 clear	Minican	$3\frac{5}{32}$	2000	4,850	19.4	96
250	120	T-4 clear	D. C. Bay.	3	2000	4,850	20.0	96
500	120	T-4 frosted	D. C. Bay.	$3\frac{7}{16}$	2000	10,100	20.2	

*Consult manufacturer's technical literature for current data, as values change frequently.
†RSC = recessed single contact.
Note: 1 in = 25.4 mm.

TABLE 26-7 Basic Data on 120 V PAR Lamps

Lamp shape	Wattage	Distributions available*	Center beam candlepower	Lumens	MOL
Standard halogen					
PAR 16	50	FL	640	400	2.88
	60	NSP, NFL	5000, 3100	650	2.88
	75	NSP, NFL	7500, 1900	900	2.88
PAR 20	35	NSP, NFL, WFL	3000, 800, 500	360	3.13
	50	NSP, NFL, WFL	4600, 1200, 900	550	3.13
PAR 30	50	NSP, NFL, FL	8800, 2300, 1300	900	3.62
	60	NSP, NFL, FL	12000, 2775, 1550	860	3.62
	75	NSP, NFL, FL	15400, 4000, 2100	1130	3.62
PAR38	45	SP, WSP, FL, WFL	1500, 700	560	5.31
	50	SP, FL	10500, 1850	650	5.31
	60	SP, WSP, NFL, FL	16000, 10500, 3700, 2500	850	5.31
	75	SP, WSP, FL, WFL	19200, 12300, 3150, 1300	1060	5.31
	90	SP, WSP, NFL, FL, WFL	19000, 14300, 4700, 3500, 1600	3110	5.31
	100	SP, FL	22000, 4000	1500	5.31
	120	SP, NFL, FL, WFL	22500, 7700, 4600, 2000	1800	5.31
	250	SP, FL	46500, 9000	3600	5.31
PAR 56	500	NSP, MFL, WFL†	78500, 40000, 19500	8800	5
PAR 64	1000	NSP, MFL, WFL‡	135000, 82000, 23000	19400	6
Halogen IR lamps					
PAR 30 IR	40	NSP, NFL, FL	8800, 2300	680	3.62
	50	NSP, NFL, FL	13000, 2900, 1400	900	3.62
PAR 38 IR	50	SP, NFL	4000, 3000	850	5.31
	55	SP, FL	14000, 2500	800	5.31
	60	SP, WSP, NFL, FL	20000, 12000, 5000, 3600	1110	5.31
	80	SP, FL	25000, 5500	1500	5.31
	100	FL, NFL, FL	29000, 6300, 3400	2070	5.31

*Beam spread for distributions are: VNSP 5°, NSP 9°, SP 9°, WSP 12°, NFL 25°, FL 30°, WFL 40 or 50°.
†8 × 15, 11 × 30, 20 × 45 degree beam spreads.
‡8 × 20, 10 × 30, 20 × 60 degree beam spreads.

design, lamp burning position, maximum wattage, limits on bulb and base temperatures, screens to protect people and surroundings, etc.

Small Tungsten-Halogen Lamps. Families of 13/8- and 2- in-diameter, 12-V models having internal multifaceted reflectors, providing a range of beam spreads for accent and display lighting, are available. Dichroic reflector coatings reduce approximately two-thirds of the heat in the beam by emitting infrared energy to the rear, thus decreasing fading of color-perishable items that the illuminated. Small changes in applied voltage have large effects on lamp life, and rapid on-off operation will shorten life. Dimmers recommended for the inductive loads of the step-down transformers should be used instead of incandescent-type dimmers. Blackening of the tungsten-halogen capsule which may result from dimming can be cleared up by full 12-V operation.

Rough and Vibration Service. These lamps are for use where lamps are subjected to shock and vibration while in use. filament construction differs. Rough-service lamps are available from 25 to 200 W, while those for vibration service range from 40 to 150 W.

Decorative Lamps. Incandescent lamps in many bulb shapes, bases, and wattages are available for a variety of decorative and architectural lighting applications. Some have specific requirements regarding burning position, shielding from moisture, etc., as covered in technical literature of the manufacturer.

26.4 FLUORESCENT LAMPS

Fluorescent lamps are low-pressure mercury electric-discharge lamps in which a phosphor coating transforms of the ultraviolet energy generated by the discharge arc into light. The major parts of a fluorescent lamp (hot-cathode type) are the bulb (tube), electrodes, fill gas, phosphor coating, and bases, as shown in Fig. 26-8. When the proper voltage is applied across the ends of the lamp, an arc is produced by current flowing between the electrodes through the fill gas (mercury vapor). This discharge generates some visible radiation, but mostly ultraviolet at 253.7 nm, which in turn excites the phosphor coating to emit light.

Fluorescent lamps are available commercially principally in four distinct types, depending upon their operating circuits: (1) hot-cathode, preheat-starting; (2) hot-cathode, instant-starting; (3) hot-cathode rapid-start; and (4) cold-cathode.

Bulb. Fluorescent lamp bulbs are basically tubular of small cross-sectional diameter. The bulb is available in straight, U-shaped, and circular configurations in bulb diameters from1/4 to 21/8 in. In straight lengths, they range from 6 to 96 in (nominal). Shorter lamps, such as the 22-, 34-, and 46-in T-5 lamps, can simplify the design of luminaires for 600-and 1200-mm module ceiling systems. Circular (circline) lamps have nominal overall diameters from $6^1/_2$ to 16 in. U-shaped lamps are 24 in (hot-cathode) and 45 in (cold-cathode) in nominal overall length. Fluorescent lamps are designated by a letter indicating the tube cross section shape and a number indicating the diameter in eighths of an inch. A T-8 lamp has a tubular bulb of 1 in diameter. Smaller diameter lamps and lower height ballasts can result in "thinner" luminaires. See Table 26-8 for a collection of typical fluorescent lamp size and wattages.

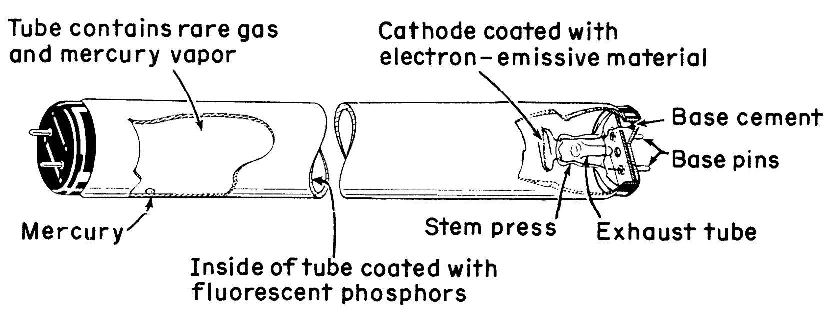


FIGURE 26-8 Cutaway view of fluorescent hot-cathode preheat-starting lamp.

TABLE 26-8 Typical Fluorescent Lamps[a]

Lamp Wattage, W	Nominal length, in[b]	Bulb	Base	Rated life, h[c]	CRI[d]	Initial lumens[e]
Preheat start—requires separate starter or starting switch						
4	6	T-5	Miniature bipin	6,000	62	135
6	8.6	T-2	Axial	10,000	82	310
6	9	T-5	Miniature bipin	7,500	62	295
8	12.6	T-2	Axial	10,000	82	500
8	12	T-5	Miniature bipin	7,500	62	400
11	16.6	T-2	Axial	10,000	82	680
13	20.6	T-2	Axial	10,000	82	860
13	21	T-5	Miniature bipin	7,500	62	850
13	12	T-8	Medium bipin	7,500	62	565
14	15	T-8	Medium bipin	7,500	62	685
14	15	T-12	Medium bipin	9,000	62	650
15	18	T-8	Medium bipin	9,000	62	825
15	18	T-12	Medium bipin	9,000	62	760
20	24	T-12	Medium bipin	9,000	62	1,200
30	36	T-8	Medium bipin	7,500	62	2,175
82[f]	60	T-17	Mogul bipin	9,000	62	5,750
90	60	T-17	Mogul bipin	9,000	62	6,000
Rapid start—straight lamps						
14	22	T-5	Miniature bipin	20,000	85	1,350
17	24	T-8	Medium bipin	20,000	75	1,325
21	34	T-5	Miniature bipin	20,000	85	2,100
25	36	T-8	Medium bipin	20,000	75	2,080
25[f]	36	T-12	Medium bipin	18,000	62	1,925
28	46	T-5	Miniature bipin	20,000	85	2,900
25[k]	48	T-12	Medium bipin	12,000	62	1,860
28[f,g]	48	T-12	Medium bipin	18,000	49	2,475
30	36	T-12	Medium bipin	18,000	62	2,275
32	48	T-8	Medium bipin	20,000	75	2,850
32	48	T-8	Medium bipin	20,000	80	2,950
32[f]	48	T-12	Medium bipin	15,000	62	2,525
32[f]	48	T-12	Medium bipin	15,000	49	2,700
34[f]	48	T-12	Medium bipin	20,000	62	2,650
34[f]	48	T-12	Medium bipin	20,000	72	2,750
34[f]	48	T-12	Medium bipin	20,000	49	2,825
35	58	T-5	Miniature bipin	20,000	85	3,650
35HO	24	T-12	Recessed DC	9,000	62	1,620
40	48	T-12	Medium bipin	24,000	73	3,300
40	48	T-12	Medium bipin	24,000	80	3,400
40	60	T-8	Medium bipin	20,000	75	3,600
45HO	36	T-12	Recessed DC	9,000	62	2,800
60HO	48	T-12	Recessed DC	12,000	62	4,050
85HO	72	T-12	Recessed DC	12,000	62	6,350
95HO	96	T-12	Recessed DC	12,000	62	8,000
110HO	48	T-12	Recessed DC	10,000	62	6,200
110HO	96	T-12	Recessed DC	12,000[i]	72	9,200
185HO	96	T-12	Recessed DC	9,000[i]	62	12,500
215HO	96	T-12	Recessed DC	10,000[i]	62	13,500
215PG	96	PG-17	Recessed DC	12,000[i]	62	15,000

TABLE 26-8 Typical Fluorescent Lamps[a] (*Continued*)

Lamp Wattage, W	Nominal length, in[b]	Bulb	Base	Rated life, h[c]	CRI[d]	Initial lumens[e]
			Rapid start—U-shaped lamps			
31U/1$^5/_8$	22$^1/_2$	T-8	Medium bipin	20,000	82	2,725
32U/6	22$^1/_2$	T-8	Medium bipin	20,000	75	2,700
35U/3[f]	22$^1/_2$	T-12	Medium bipin	18,000	62	2,350
34U/6[f]	22$^1/_2$	T-12	Medium bipin	18,000	70	2730
40U/3	22$^1/_2$	T-12	Medium bipin	18,000	70	2925
40U/6	22$^1/_2$	T-12	Medium bipin	18,000	70	3050

[a]Check manufacturers' technical literature for current data, as values change frequently.
[b]Includes lamp and two standard lampholders, except RS T-5 lamps.
[c]lamp burning hours to median life expectancy, when operated 3 h per start. More frequent starting reduces life and less frequent starting increases life. Some energy-saving lamps on single-lamp ballasts may have shorter life. Ballast strongly affects lamp life.
[d]Color rendering index (CRI); rates ability to render color of objects on a scale of 0 to 100. Numerical values should be compared only for lamps of the same color temperature.
[e]After 100 burning hours. Consult ballast or luminaire manufacturer for appropriate multiplier (*ballast factor*). Some lamp catalogs also give mean or design lumens.
[f]Use in 60°F or higher ambients, protect lamp surfaces from strong air drafts, and check manufacturer about operation on dimming, reduced current, or cathode cutout systems.
[g]Based on 12 h per start.

U-shaped models employing $^1/_2$- or $^5/_8$- in-diameter bulbs with little separation between the legs are called "compact fluorescent." in lengths below 9 in, they can have one or more twin tubes. Family designations vary. Subminiature hot- and cold- cathode tubular types are 0.25 and 0.266 in in diameter, respectively, and with lengths from 4 to 20 in.

Electrodes. There are two electrodes in each fluorescent lamp, one at each end, designed to operate as either "hot" or "cold" electrodes (or cathodes).

Hot-cathode lamps contain electrodes which are usually coiled-coil (or triple-coiled) tungsten filaments coated with one or more of the alkaline-earth oxides. By suitable circuit arrangements these cathodes can be heated to an electron-emitting temperature before the arc strikes, or they may be required to act momentarily as cold cathodes until they are heated by bombardment after the lamps ha started. Lamps using these cathodes may be designed to carry currents of 1 to 2 A with low-voltage drop (10 to 12 V) at the electrodes. Some energy-saving types of rapid-start ballasts have disconnect elements to discontinue cathode heating after the lamp starts. The power saved is approximately 3 W per lamp. Metal shields can be used to minimize end darkening, improving lamp lumen maintenance.

Cold-cathode lamps are those that use electrodes of tubular form of iron or nickel which may be coated on their inside surfaces with electron-emitting materials. These cathodes operate at temperatures which limit the lamps to low-current densities. The electrode drop in these lamps is relatively high (over 50 V), but they are not subject to short life as a result of frequent instant starting.

Fill Gas. Droplets of liquid mercury are present in the fluorescent lamp and vaporize to a very low pressure during lamp operation. Argon is added to assist ignition of the discharge in standard lamps, while energy-saving types have an argan-krypton mixture. Certain other types use a combination of argon and neon or argon, neon, and zenon.

Phosphors. The chemical composition of the phosphor coating on the bulb interior surface determines the color of the light produced and, in part, lamp efficacy. Those lamps with phosphors producing good overall color rendering are generally of higher efficacy. Figure 26-9 shows typical spectral power distributions for a variety of different phosphor compositions.

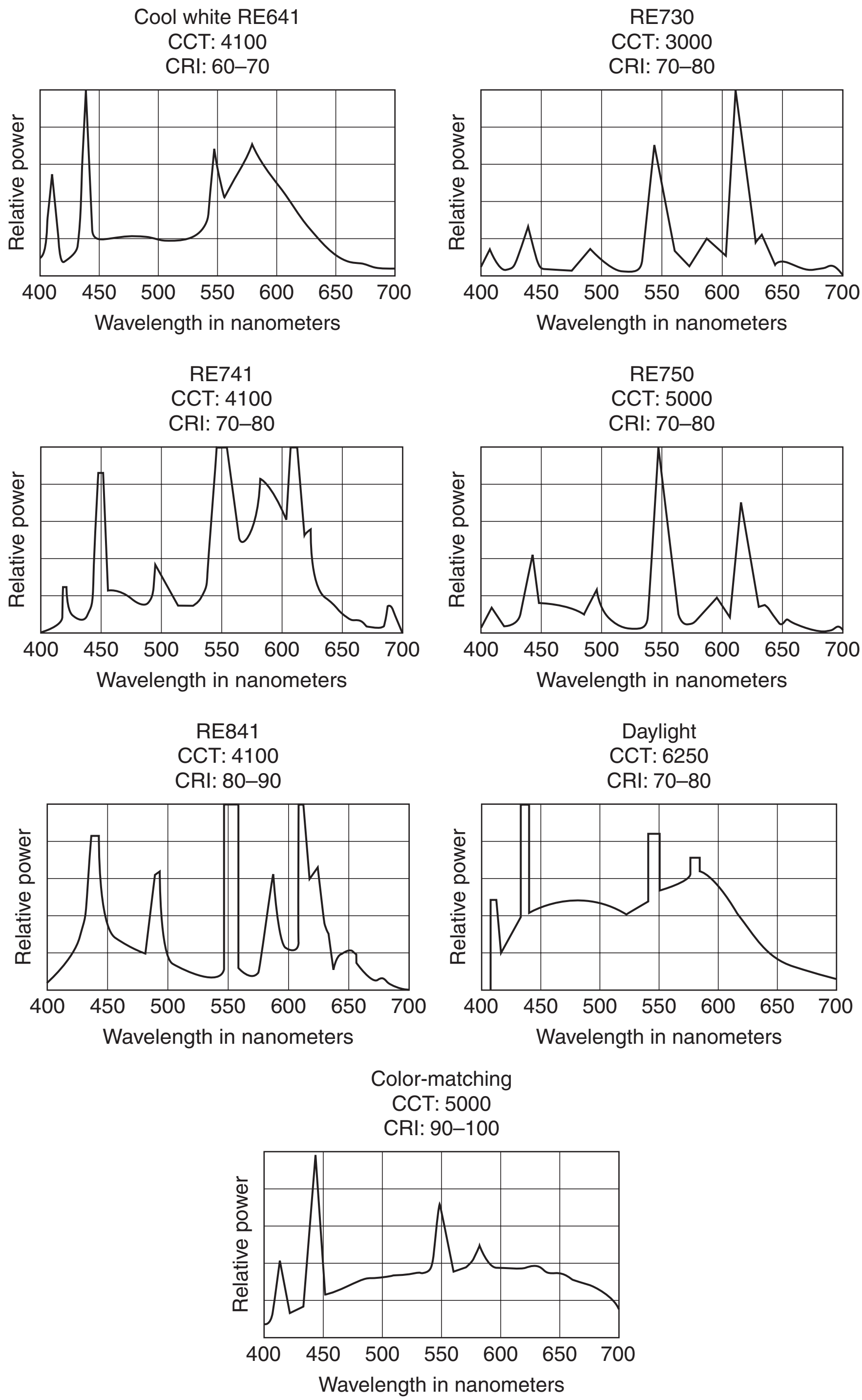


FIGURE 26-9 Spectral distribution curves for typical fluorescent lamps. (*Reprinted from the IESNA Lighting Handbook 9th ed. with permission from the IESNA.*)

The bulbs of some fluorescent lamps have a single, thick inner coat of conventional “halophosphor.” Adding a thin coat of more expensive, rare-earth triphosphors can provide an improved color rendering index (CRI) and increase the efficacy. When a double coat of the triphosphors is used, CRIs are 80 to 90, while retaining the higher levels of lumens per watt. Triphosphor coatings are standard on certain families of fluorescent lamps—check manufactures’ technical literature for current information, and for designations employed with superior-color lamps. Special phosphors are also used in fluorescent lamps designed for plant growth and for black-light effects.

Bases. Lamps designed for instant-start operation generally have a base at each end with a single pin connection. (In some cases instant-start lamps may have two pins at each end electrically connected.) Lamps for preheat or rapid-start operation also have a base at each end, but with two pins (connections) in each. Some manufacturers use a green base finish or print to identify fluorescent lamps which have less mercury and/or pass the toxicity characteristic leaching procedure (TCLP), and therefore are classified as nonhazardous waste in many states. Rapid-start high-output lamps have recessed double-contact bases, and T-2 subminiature fluorescent lamps have axial bases. The circline lamp has a single four-pin connector. Compact fluorescent lamps may have single two-pin or four-pin bases. Four-pin bases are required if the lamps are to be dimmed. See Fig. 26-10 for images of the available fluorescent lamp base types.

Compact Fluorescent Lamps. Energy-conservation activities have focused attention on the relatively low efficacy and short life of general-service 25- to 100-W incandescent lamps widely used in many residential, commercial, institutional, and industrial applications. Compact fluorescent lamps provide significantly higher efficacy and come in a variety of sizes and wattages (see Table 26-9). Shorter models can replace incandescent lamps in existing and new table and floor lamps, in recessed downlights, etc. The $10^1/_2$- and $22^1/_2$-in-long sizes are useful in 1- and 2-ft luminaires, and a three-lamp, 2 by 2 ft recessed luminaire with 40-W twin-tube lamps can achieve so-called “nondirectional” layouts without significant reduction in total luminaire output

TABLE 26-9 Typical “Compact” and Longer Twin Tube Fluorescent Lamps with Pin Bases

Watts	Generic lamp designation	Number of twin tubes	Bulb	Nominal length	Initial lumens
5	CFT5	1	T4	3.4	230
7	CFT7	1	T4	4.5	400
9	CFT9	1	T4	5.7	580
13	CFT13	1	T4	6.2	820
9	CFQ9	2	T4	4.3	560
10	CFQ10	2	T4	4.0	600
13	CFQ13	2	T4	5.2	860
18	CFQ18	2	T4	5.8	1160
26	CFQ26	2	T4	6.9	1700
13	CFTR13	3	T4	4.2	900
18	CFTR18	3	T4	4.4	1200
26	CFTR26	3	T4	5.0	1710
32	CFTR32	3	T4	5.6	2200
42	CFTR42	3	T4	6.4	3200
57	CFTR57	3	T4	7.8	4300
70	CFTR70	3	T4	9.3	5200
32	FT32	1	T4	5.8	2200
39	FT39	1	T5	16.5	2850
39	FT39	1	T5	22.5	3150
50	FT50	1	T5	22.5	4000

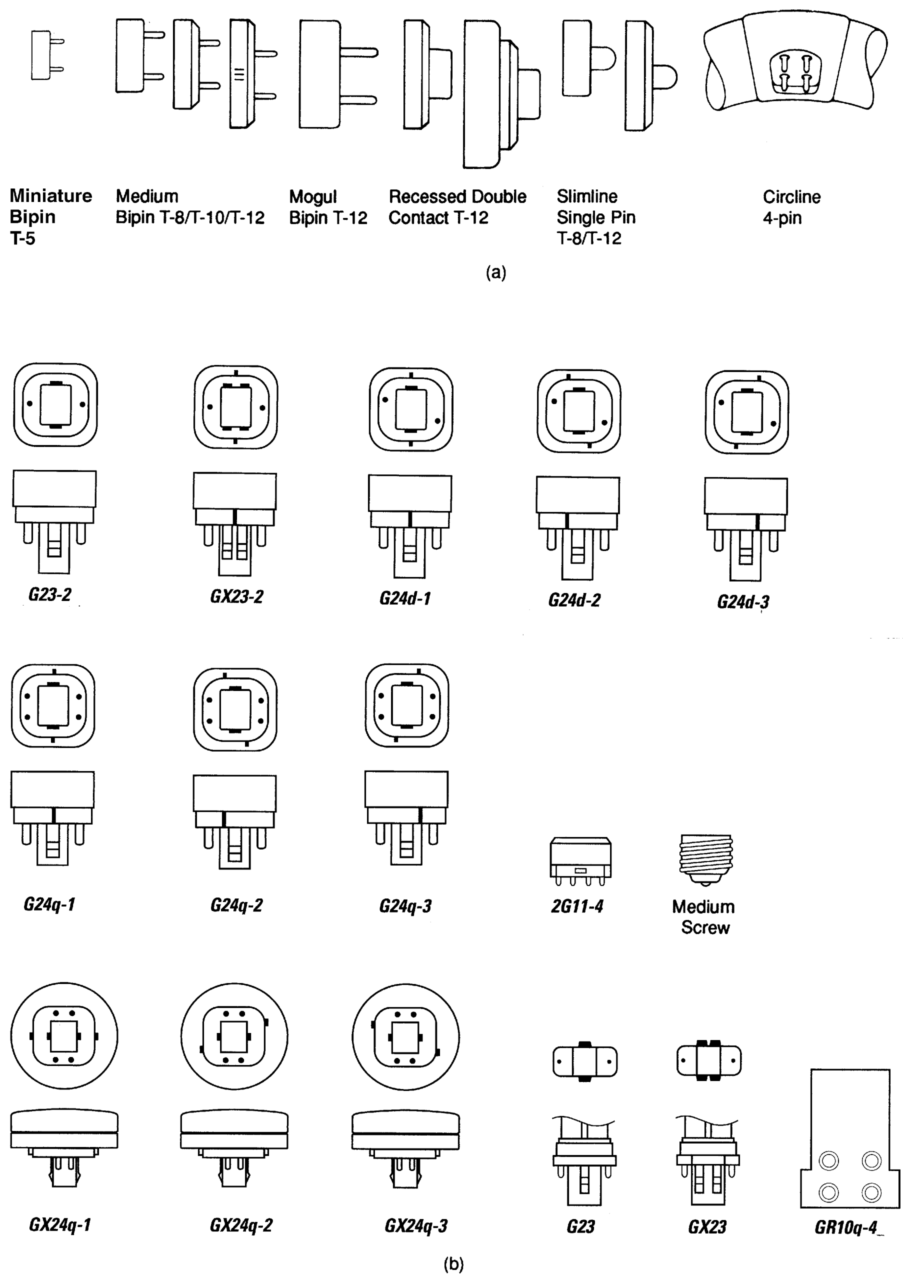


FIGURE 26-10 Base used for common types of fluorescent lamps: (*a*) regular fluorescent lamps bases, (*b*) compact florescent lamps bases.

compared with 2 × 4 ft units. Designers should check with lamp manufacturers about suitability for dimming.

Medium-screw base adapters are available for installing certain compact fluorescent lamps in 120-V as sockets. Some adapters have integral preheat ballasts, and others contain electronic components. Permanently assembled lamps, starter, and ballast units are also available with medium-screw bases.

Certain adapters can be used with bare compact fluorescent lamps, while other adapters have enclosing globes of various shapes, or reflectors to obtain directional light. Typical input wattages are 5, 9, 13, 18, and 26, considered as replacements for 25-, 40-, 60-, 75-, and 100-W incandescent lamps, respectively.

Energy Distribution. The approximate distribution of energy in a typical cool white fluorescent lamps is shown in Fig. 26-11.

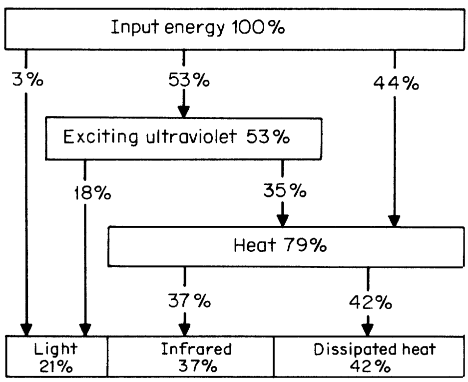


FIGURE 26-11 Energy distribution in a typical fluorescent lamp.

Performance Characteristics. Table 26-8 lists some typical fluorescent lamps for general lighting along with their physical characteristic, rated life, color-rendering index, and rated initial lumen output. Consult ballast manufacturers for input wattage data needed for energy calculations. Table 26-10. contains data for a variety of fluorescent lamp/ballast combinations. Due to the significant savings provided by electronic ballasts, most fluorescent ballasts used today in commercial luminaires are electronic.

The escalation of electronic data-processing equipment and variable-speed motors increased attention given to the harmonic content of electric power systems. Fluorescent-lamp ballasts are known to contribute to total harmonic distortion (THD). Harmonics raise the current in the neutral conductor of 3-phase, 4-wire, wye-connected power distribution systems, even though the phase loads may be reasonably balanced. Some older circuits exist where reduced neutrals are used for fluorescent lighting loads. However, full 100%

TABLE 26-10 Wattage and Efficacy Data for a Variety of Fluorescent Lamp/Ballast Combinations

Lamp designation	Lumens	Lamp wattage	Nominal length in	Ballast type	Ballast factor	Single-lamp ballast input watts	Two-lamp ballast input watts	Efficacy
F34T12		34	48	Electronic	0.9	30	59	
F34T12		34	48	Magnetic	0.91/0.89	44	77	
F40T12		40	48	Electronic	0.88/0.84	38	69	
F17T8		17	24	Electronic	0.95/0.98	19	34	
F32T8		32	48	Electronic	1.2		78	
F32T8		32	48	Electronic	0.88	30	59	
F54T5HO		54	46	Electronic	1	62/60	121/118	
F96T12		75	96	Electronic	0.88		135	
F96T12HO		110	96	Electronic	0.88		210	
F96T8		86	96	Electronic	0.88		112/110	
FT40		40	22.6	Electronic	0.88	38/37	76/73	

capacity neutral conductors have long been recommended for branch circuits serving loads consisting of more than one-half fluorescent lighting. Indeed, some electrical engineers specify cables with single, oversized neutral conductors, cables providing a separate neutral for each phase, transformers designed to handle harmonic loading, etc.

Light output for fluorescent lamps is sensitive to surrounding (ambient) air temperature as shown in Fig. 26-12. Lamp wattage changes in a similar fashion, but not as drastically at ambients below normal. Lamps operated at ambient temperatures below 60°F should be enclosed to conserve their heat. Air movement over the lamp bulb has the effect of lowered ambient temperature. Some CFLs apply a mercury amalgam that provides more stable lumen output over a wider range of temperatures and operating positions.

Fluorescent lamps generally should be operated at voltages within ± 10% of their designed operating points for best performance. Decreased life land uncertain starting may result from operation at lower voltages, and at higher voltages there is danger of overheating of the ballast as well as decreased lamp life. One exception to this is found in the series operation of cold-cathode lamps, where an adjustable voltage supply makes possible operation over a wide range of illumination levels, that is, dimmer operation such as that used in stage lighting.

Failure of a hot-cathode fluorescent lamp usually results from loss of active material from the cathode or cathodes. This loss proceeds gradually throughout the life of the lamp and is accelerated by frequent starting. Depreciation of light output is caused principally by tube blackening and is rapid (as much as 10%) during the first 100 h but very gradual from that point on. For this reason the lamps are rated commercially on the basis of the lumen output after 100 h of operation.

Fluorescent-lamp Operation. Fluorescent lamps are best adapted to operate on ac circuits with reactance ballasts. Typical operating circuits are shown in Figs. 26-13 to 26-17.

Fluorescent lamps are, to a considerable extent, dependent on the characteristics of the ballast equipment. Typical of this is the effect of variations from rated line voltage on the conditions of lamp operation. Certified ballasts made in accordance with industry specifications and periodically field-checked by an independent laboratory are available for the more commonly used fluorescent lamps. They are to be distinguished by the letters CBM on the ballast case. Thermally protected "Class P" ballasts are required for fluorescent fixtures installed indoors, except fixtures with simple reactance ballasts.

The fluorescent lamps, in itself, is inherently a high-power-factor circuit, but the reactive ballast normally used to stabilize the arc is inherently low power factor. Since in the usual circuit the voltage drop across the ballast is approximately equal to that across the lamp arc, the resulting power factor of a single-lamp reactive-ballast circuit is on the order of 50%. For many applications this low power factor is objectionable. In single-lamp ballasts, power factor correction may be obtained by means of a capacitor shunted across the line connections or, where the

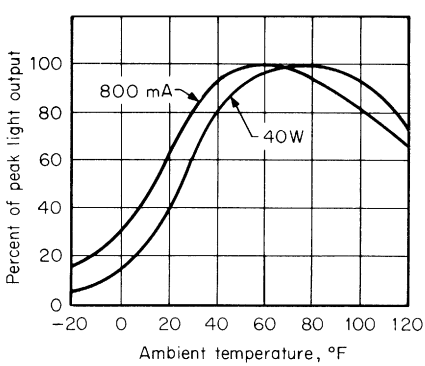


FIGURE 26-12 Effect of air temperature on light output for a typical fluorescent lamp

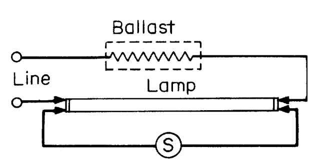


FIGURE 26-13 Single-lamp ballast for 4- to 40-W hot-cathode, preheat-starting fluorescent lamp (S = starting switch).

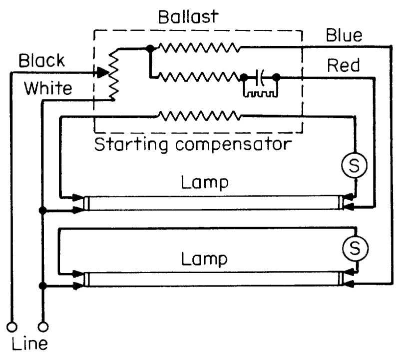


FIGURE 26-14 Two-lamp ballast circuit for 30- and 40-W hot cathode, preheat-starting fluorescent lamps, showing built-in starting compensator.

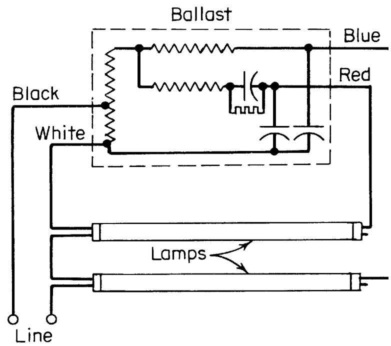


FIGURE 26-15 Two-lamp lead-lag ballast circuit for instant-starting hot-cathode lamps.

lamp requires a higher voltage, by a capacitor across the transformer secondary. The two-lamp ballast, through phase displacement of the lamp currents, or series capacitors, offers a ready means of power factor correction and is usually designed to give a circuit power factor greater than 90%.

All inductive fluorescent ballasts emit a certain amount of noise; the noise increases with the lamp current. A sound rating for ballasts has been developed by some manufactures from A (quietest) of F (noisiest). The amount of cumulative ballast noise, which is tolerable, depends on two sets of principal factors: (1) characteristics of the room and (2) characteristics of the luminaire. Electronic ballasts are generally much quieter than magnetic ballasts and provide an "A" rating.

Where direct current is available at circuit voltages comparable with the open-circuit voltages of the usual ac ballast circuits, fluorescent lamps may be operated from these sources. For such operation, resistance must be added to the usual series reactance ballast (transformer ballasts are not applicable) to limit the operating current to the designed value. This causes a marked reduction in the overall efficacy of the lamp and circuit combination over that obtained in ac operation. Under dc operation, lamps more than a few feet in length will promptly develop a concentration of the mercury vapor at the negative end of the lamps, with the result that only a fraction of the bulb will give

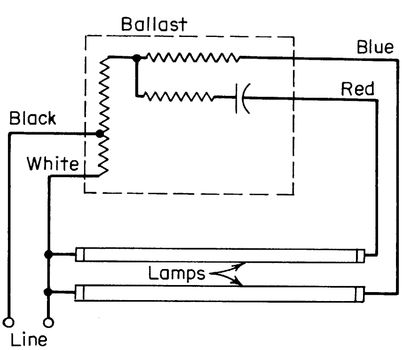


FIGURE 26-16 Two-lamp lead-lag ballast circuit for multiple operation of cold-cathode lamps.

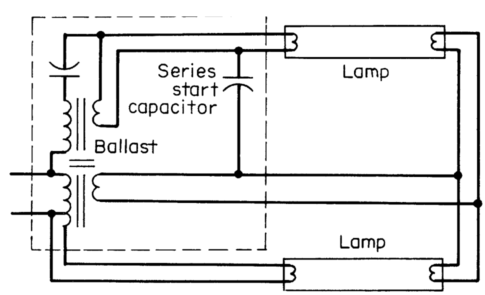


FIGURE 26-17 Two-lamp series rapid-start high-power-factor circuit.

off light. This condition can be overcome through a periodic (about once in 4 h) reversal of the polarity of the lines feeding the lamps. The life of lamps is likely to be shorter on dc.

Dimming. For dimming hot-cathode fluorescent lamps, a number of different arrangements are available. For smooth operation, the lamps on any circuit should be made by the same manufacturer, at the same time, in the same color, and of the same age in use. Group replacement is the most satisfactory procedure. Lamps should be operated free from drafts at 50 to 80°F and should be seasoned 100 h at full brightness prior to dimming. Certain energy-saving lamps are not recommended for dimming applications. Special dimming ballasts are typically required.

Dimming ballasts are available to reduce light output of T-8 and T-12 rapid-start lamps to 1%, 5%, or 10% of full lighting output. Most dimming ballasts are electronic ballasts. Standard T5 lamps can only be dimmed to 5% while T5HO can be dimmed to 1%. Compact fluorescent lamps can be dimmed to either 1% or 5%. Dimming ballasts offer slightly lower lumens per watt than nondimming electronic ballasts. One manufacturer states that dimming from 100% to 1% and to 10% is perceived as 10% and 32% of full brightness, respectively. Various sliding and other types of wall box and wireless dimmer controls are available for dimming control, or in connection with occupant and daylight sensors for automatic energy-conservation systems. Ballasts can be controlled via analog or digital signals, and the controller must be configured to operate the type of ballasts being used. Compatibility with emergency-lighting ballasts should also be explored.

Electronic Ballasts. The use of solid-state electronic elements instead of magnetic components can increase lamp efficacy by higher-frequency operation of lamps, and can reduce input power. This has made electronic ballasts the standard for today's fluorescent lighting systems. Some use a control chip that results in constant light output and energy consumption over a significant range of line voltages. Electronic ballasts are lighter in weight, operate cooler, and can be designed for rapid-and instant-start lamps to meet federal efficacy and FCC EMI/RFI standards. Some models permit dimming of fluorescent lamps, and can be used with appropriate sensors to compensate for changes in daylight illuminance levels. Electronic ballasts cost more than electromagnetic types, but often provide swift payback for the lighting-hours usage and kilowatt-hour rates typical of commercial, institutional, and industrial applications.

Table 26-11 shows the energy-saving potential of electronic ballasts. The annual operating cost values can be adjusted for burning hours other than 4000 h per year, and for kilowatt-hour rates lower or higher than 7 cents. Note that the combination of T-8 lamps and electronic ballasts gives higher system efficacy than other lamp-ballast combinations in the table. Ballast factor is the lumen output of lamps operated on commercial ballasts divided by the lumen output of those lamps operated on a reference ballast.

CBM electronic ballasts have ballast factors of 0.85 or higher, compared with 0.925 to 0.95 for CBM magnetic ballasts. However, fluorescent lamps operate at lower bulb-wall temperatures on electronic ballasts, so a ballast factor of 0.85 will result in approximately the same light output when within enclosed luminaires, compared to magnetic ballasts.

Both electronic and magnetic ballasts generate harmonics in the line current. For electronic types, most modern fluorescent ballasts limit the total harmonic distortion to under 20% or under 10%, neither of which should create problematic current in the neutral conductor. Engineers should check the current ANSI and IEEE 519/587 standards. Reported measurements of compact fluorescent lamps and diode devices for use in incandescent-lamp sockets showed power factors in the 47% to 67% range, and total harmonic distortion (THD) greater than 100%.[1] Although electronic ballasts generally have lower THDs than do magnetic ballasts, check compliance with the requirement of most electric utilities that the THD of electronic ballasts be less then 20%.

Many electronic components are employed in assembling electronic ballasts. Their specific designs are considered proprietary, but Fig. 26-18 gives the block diagram for basic electronic ballasts for fluorescent lamps.*

*Provided by J. N. Lester, Osram/Sylvania, Inc, Beverly, Mass.

TABLE 26–11 Luminous and Electrical Characteristics of a Three-Lamp VDT-Type Recessed Luminaire Using Various Lamp and Ballast Combinations

			Luminaire			Comparative			
Lamp	Ballast	Ballast factor	Efficiency	Watts input	Lumen output	Watts input, %	Lumen output, %	System efficacy, %	Annual* operating cost, $
F40-T12ES 2650 lm	Magnetic ESB	0.88	0.674	117	5355	Base 100	Base 100	Base 100	32.76
F32-T8 2900 lm	Magnetic for F32-T8	0.94	0.656	111	5362	95	100	120	31.08
F40-T12ES 2650	Electronic	0.82	0.668	88	4352	75	81	123	24.64
F32-T8 2900 lm	Electronic	0.93	0.719	90	6253	77	117	160	25.20

*Based on 4000 h at 7¢/kWh. Information furnished by Thomas Industries, Inc. Check manufacturers' technical literature, as data change frequently.

Formulas used:

$$\text{Annual operating cost} = \text{W} \times \text{h / year} \times \frac{\text{cost/kWh}}{1000}$$

$$\text{System efficacy} = \frac{\text{basic fixture efficiency} \times \text{lamp lumens} \times \text{ballast factor}}{\text{input watts}}$$

$$\text{Simple payback (in years)} = \frac{\text{additional cost}}{\text{savings / year}}$$

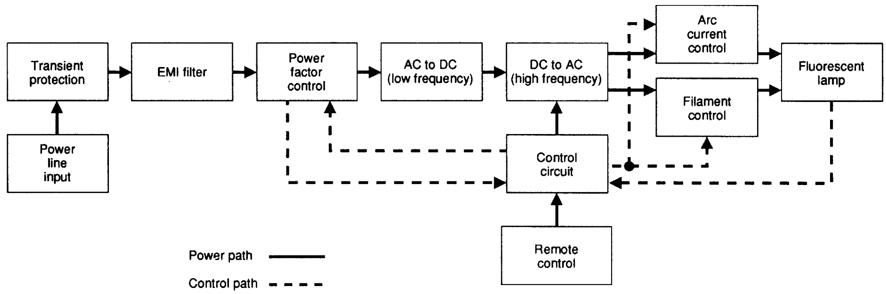


FIGURE 26-18 Basic block diagram for electronic fluorescent lamp ballasts. (*Courtesy by Osram/Sylvania Inc.*)

Most ballasts for T-2, T-4 and T-5 fluorescent lamps sense deactivated-cathode "end-of lamp life" conditions when one or more of the cathodes are depleted. This avoids a potentially hazardous situation when those small-diameter lamps are used with electronic ballasts, which could continue operating a bad-cathode lamp, perhaps causing the glass at that end to melt or crack. Some electronic ballasts for T-8 rapid-start lamps are expected to also incorporate end-of-life shutdown circuits.

Electronic ballasts are classified as rapid-start, instant-start, and programmed start. The latter designation is used for ballasts that contain a microprocessor programmed to preheat the cathodes and increase the starting voltage at an optimum temperature, and to sense end-of-lamp life and disconnect the ballast until the lamp is replaced.

Electronic instant-start ballasts give shorter lamp life than do rapid-start and programmed-start ballasts, but are lower in input wattage. The savings in kilowatt-hours often makes instant-start electronic ballasts the economic choice, unless occupancy sensors (or other causes of frequent lamp cycling) turn the lighting off and on 7 or more times each day. Users should also check with electronic ballast manufacture about level of line inrush current on starting.

26.5 HIGH-INTENSITY DISCHARGE LAMPS

High-Intensity Discharge (HID). This term denotes a general group of lamps consisting of mercury, metal halide, and high-pressure sodium lamps. A mercury lamp is an electric discharge lamp in which the major portion of the radiation is produced by the excitation of mercury atoms. A metal halide lamps is an electric discharge lamp in which the light is produced by the radiation from an excited mixture of a metallic vapor (mercury) and the products of the dissociation of halides (for example, halides of thallium, indium, sodium). A high-pressure sodium lamp is an electric discharge lamp in which the radiation is produced by the excitation of sodium vapor in which the partial pressure of the vapor during operation is of the order of 10^4 N/m^2.

Lamp Construction and Designation. HID lamps consist of a cylindrical transparent or translucent arc tube which confines the electric discharge and the associated gases. That tube is further enclosed in a glass bulb or outer jacket to exclude air to prevent oxidation of the metal parts and to stabilize operating temperatures and significantly reduce ultraviolet radiation emitted by the excitation of the vapors. The mount structure of many HID lamps is anchored to the "dimple top" of the outer glass bulb, assuring greater structural integrity and more accurate alignment of the arc tube. The construction of a typical mercury lamp is shown in Fig. 26-19. The basic elements are the arc tube, fabricated from fused silica and filled with a drop of mercury and a rare gas at low pressure; the electrodes; and the outer envelope, which may or may not have a phosphor coating on the

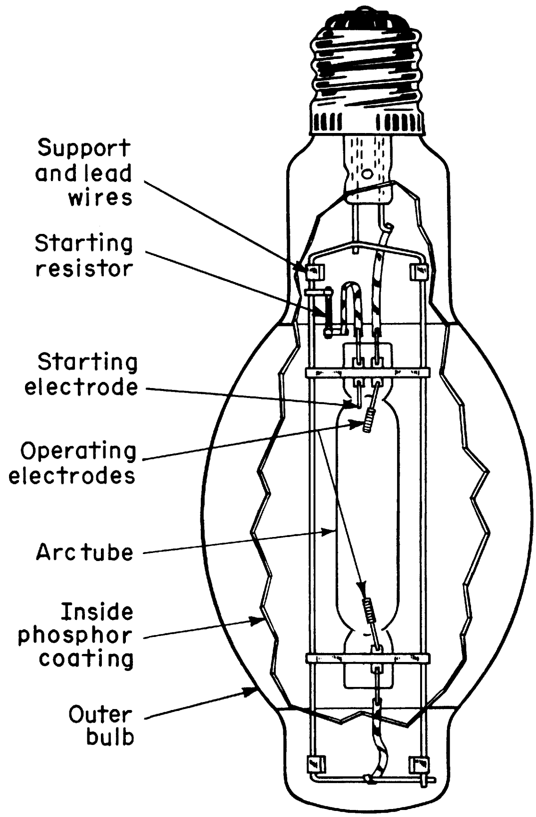


FIGURE 26-19 A 400-W phosphor-coated mercury lamp. Lamps of other sizes are constructed similarly.

interior for improved color rendering. Mercury lamps, due to their relative inefficiency and generally poor color rendering, are rarely, if ever, applied in today's lighting designs.

Metal halide lamps are very similar in construction to the mercury lamp, the major difference being the addition of a metal halide in the arc tube (see Fig. 26-20). The outer bulb may or may not have an inner phosphor coating to improve color rendition, and for lower, more uniform lamp brightness. Special metal halide lamps, with aluminized reflective coatings on the top of the bulb, can reduce glare and help minimize light trespass. There are PAR-and R-bulb mercury and metal halide lamps, xenon metal halide for fiber optic systems, and iodine metal halide to simulate natural daylight.

Metal Halide lamps have historically been susceptible to color maintenance problems due to shifts in their color over time. Recent developments aimed to minimize the occurrence of this problem include special rounded-shape arc tubes, pulse start ignitor technology, and the use of ceramic arc tubes.

The construction of a typical high-pressure sodium lamp is shown in Fig. 26-21. The basic components are the arc tube of translucent polycrystalline or single-crystal alumina, filled with sodium, mercury, and a rare gas (xenon); electrodes; and an outer borosilicate glass envelope. This outer bulb is either clear or contains an inner diffuse coating for more uniform lamp brightness. HPS lamps with twin arc tubes provide quick-restarting when momentary power failures occur. Rated at 40,000-h life, they also are valuable for difficult-to-reach locations.

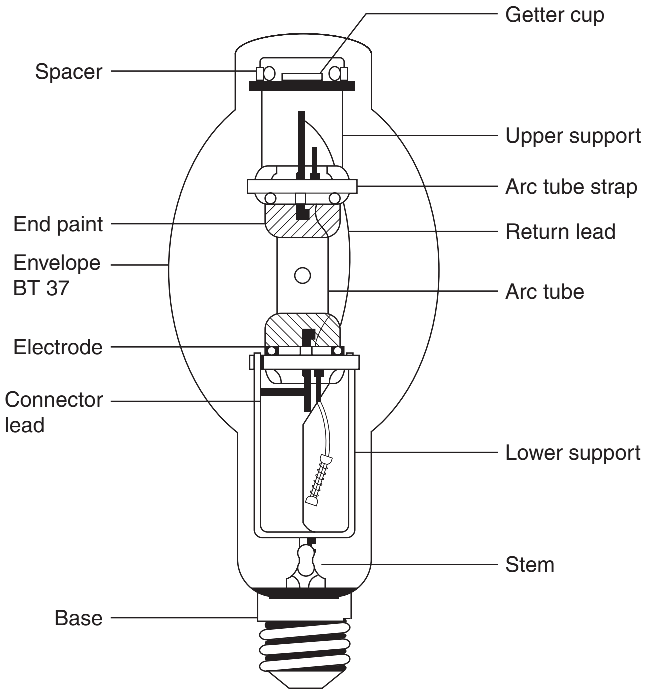


FIGURE 26-20 Construction of a standard metal halide lamp. (*Courtesy of Osram Sylvana.*)

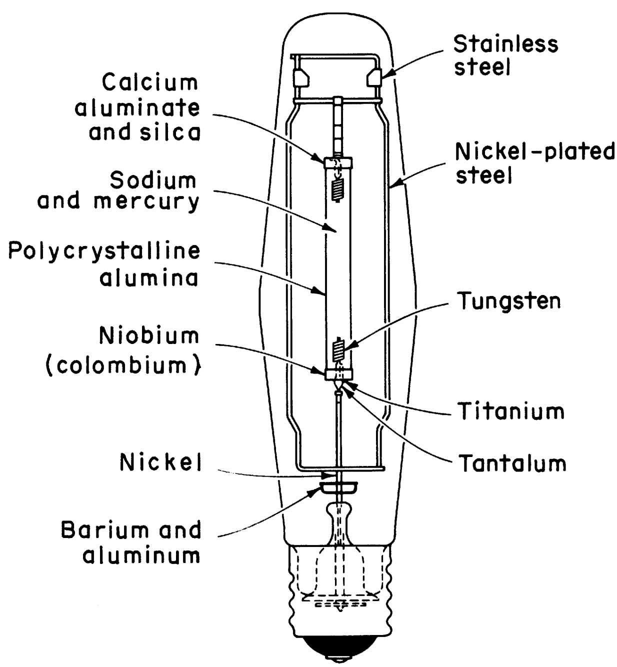


FIGURE 26-21 Construction of a typical high-pressure sodium lamp.

A lamp designation system developed by the American National standards Institute (ANSI) is currently in use.[1] It consists of five groups of letters or numbers: first a letter indicating the type of lamp (H, mercury; M, metal halide; S, high-pressure sodium), followed by an arbitrary number designating electrical characteristics (which relates to the type of ballast required), followed by two arbitrary letters which describe the physical characteristics, then the lamp nominal wattage, and finally letters indicating the phosphor color. An example for a 175-W metal halide lamp would be M57/C/175/U/MED.

Lamp Characteristics. Light output from each of the three types of HID lamps has its own color appearance (chromaticity), and the spectral power distributions vary as shown as in Fig. 26-22. Table 26-12 lists the radiated energy of typical 400-W HID lamps.

Performance Characteristics. Table 26-13 lists some typical metal halide and high-pressure sodium HID lamps for general lighting along with their light output (reference initial and mean lumens). The basis for published data may vary with manufacturer. For a qualitative comparison of HID lamps with incandescent and fluorescent lamps, see Table 26-14.

Many mercury and high-pressure sodium lamps, and some metal halide lamps, can be operated in any position. Other metal halide lamps have restricted narrow ranges of acceptable burning positions, outside of which light output and rated life may be adversely affected. Some metal halide lamps must be used only in enclosed luminaires, or may operate at higher temperatures which would exceed the temperature rating of explosion-proof or other hazardous-area luminaires. Certain metal halide lamps have compact outer bulbs and bases at each end, for smaller display-, sports- and flood-lighting luminaires. Some HID lamps have sufficient ultraviolet output to produce skin burn and/or eye injury, thus requiring specialized luminaires equipped with safety interlocks.

A new family of pulse-start metal halide lamps resulted from studies showing that improved lumen maintenance and color stability of metal halide lamps could result form (1) reducing the

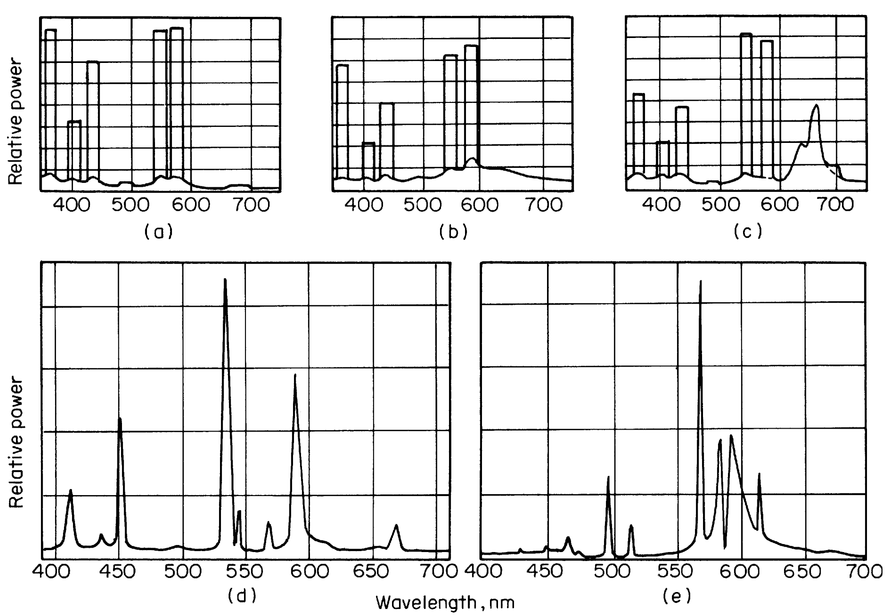


FIGURE 26-22 Spectral distribution curves for typical HID lamps: (*a*) clear mercury, (*b*) phosphor-coated mercury, (*c*) improved phosphor-coated mercury, (*d*) sodium-thallium-indium iodide metal halide, (*e*) high-pressure sodium.

TABLE 26-12 Energy Output for Some HID Lamps

Type of energy	400-W Mercury	400-W metal halide	400-W high Pressure sodium
Light	14.6%	20.6%	25.5%
Infrared	46.4	31.9	37.2
Ultraviolet	1.9	2.7	0.2
Conduction-convection	27.0	31.1	22.2
Ballast	10.1	13.7	14.9

sputtering of tungsten from the electrodes and (2) shortening the starting time.[2] Pulse start metal halide lamps require an ignitor with the proper ballast, and provide increased light output, longer life, quicker warmup, and faster hot restrike as additional benefits.

Another relatively new technology is that of ceramic arc-tube metal-halide lamps. These lamps are generally available in the lower wattages, some in PAR shapes, and provide improved color consistency and very good color rendering. The clear arc tube is replaced with a short translucent ceramic arc tube of similar material to that used in a high-pressure sodium lamp.

HID Lamp Operation. The practical limit of an HID lamp's current-carrying capacity is how high a temperature its enclosing tube can withstand without rupturing. By connecting an impedance in series with the lamp, the current is controlled. In most lamps about one-half the supply voltage is absorbed by a series ballasting device. A variety of ballasts are available for operating lamps, singly or in pairs. Single lamp ballasts may have low (0.50 minimum) or high (0.90 minimum) power factor, 2 lamp ballasts have inherently high power factor. The simplest lamp ballast is the reactor-type used in series with the lamp when line voltage is sufficient for reliable starting. This is not recommended where line-voltage fluctuations exceed 5%. A reactor-type ballast can be used when the line voltage is approximately twice the rated lamp voltage. The autotransformer-type ballast is used on circuits where the line voltage must be changed to suit the lamp requirements. Constant-wattage types of the autotransformer or isolated-secondary design are widely used because of better regulation, low line starting currents, and lower dropout voltage. They are also called *stabilized* or *regulated*. Heavier wiring, oversized circuit breakers, and time-delay relays that may be required by the relatively high starting currents of non-CW and non-CWA ballasts are eliminated with stabilizing ballasts, as the starting current is less than the rated operating current.

One of the limitations of the HID lamp is the effect of power-supply interruptions. In the event of a power interruption or voltage dip lasting for more than 1 cycle, HID lamps extinguish and do not restart for several minutes. The exact magnitude of the voltage drop to cause this condition depends on the ballast design. Regulator ballasts withstand a greater drop than other types. The delay in lamp restarting is caused by the high pressure that develops in the arc tube during operation. The open-circuit voltage of standard ballasts is not sufficient to restart the lamp until the lamp cools and the pressure decreases. In installations where this characteristic might be a safety hazard, the use of a few incandescent or fluorescent luminaires along with the HID units assures emergency illumination until the HID lamps restart. Tungsten-halogen auxiliaries are available for HID industrial luminaires to provide standby illumination in the event of momentary power failure. For indoor and outdoor sports lighting, and other applications where instant restrike is preferable, special ignitor are available, as are special instant-restrike high-pressure sodium lamps. Metal halide lamps with a wire lead (at the end opposite the base) are used with auxiliary ignitors to achieve instant restrike. For aisle lighting in warehouses and other interior and exterior situations where illumination levels for accurate seeing are not needed all of the time, high/low electrical components, combined with occupancy detectors, transmitters, and luminaire-mounted receiver scan provide major energy saving by reducing input wattage 50% to 70% during intervals when the spaces are unoccupied.

TABLE 26-13 Metal Halide and High-Pressure Sodium Hid Lamps[a]

Watts	ANSI ballast type	Bulb	Bulb finish	Base	Maximum overall length, in	Vertical burning rated life, h	Initial	Mean	CRI
Pulse-start metal halide—vertically base up +/–15°									
32	M100	ED-17	Coated	Med	5.43	10,000	2,400	1.700	70
50	M100	ED-17 or BD-17	Clear	Med	5.43	10,000	3,900	3,200	70
70	M98	ED-17 or BD-17	Clear	Med	5.43	12,000	5,500	3,500	70
100	M90	ED-17 or BD-17	Clear	Med	5.43	15,000	9,000	6,200	70
150	M102	ED-17 or BD-17	Clear	Med	5.43	15,000	12,500	8,600	70
175	M137	BD-17	Clear	Mog	5.75	15,000	17,000	12,500	65
		ED23.5			7.5				75
250	M138	ED-28	Clear	Mog	8.31	20,000	23,000	17,000	65
350	M131	ED-37	Clear	Mog	11.5	20,000	37,000	27,500	65
400	M135	ED-37	Clear	Mog	11.5	20,000	44,000	28,500	65
1000	M141	BT-37	Clear	Mog					
Ceramic MH									
39	M130	T4.5 T6	Clear	G	-	9,000	3,300	2,640	82
50	M110	ED-17	Clear	Med	5.43	10,000	4,100	2,750	85
70	M98	ED	Clear	Med	5.43	16,000	6,200	4,585	85
100	M90	ED	Clear	Med	5.43	16,000	9,500	7,125	85
150	M102	ED	Clear	Med	5.43	10,000	14,000	10,500	85
Ceramic MH PAR lamps					CBCP				
39	M130	PAR20	SP 10	Med	22,000	9,000	2,000	-	85
			FL 30		5,000				
39	M130	PAR30L	SP 10	Med	39,600	9,000	2,000	-	85
			FL 30		7,400				
70	M98	PAR30L N	SP 12	Med	46,000	9,000	4,400	-	85
			FL 30		13,200				
70	M98	PAR38	SP 15	Med	40,000	10,000	4,300	-	85
			FL 25		16,000				
			WFL6 5		3,500				
100	M90	PAR38	SP 15	Med Skt	45,000	12,000	6,500	-	85
			FL 25		25,000				
			WFL6 5		12,000				
Metal halide—except as noted, vertically base up or base down +/–15°									
175	M57	BD-17 ED-28	Clear	Mog	8.25	10,000	13,500	9,100	
250	M58	ED-28	Clear	Mog	8.25	10,000	20,500	13,500	
400	M59	ED-28 ED-37	Clear	Mog	11.3	20,000	36,000	23,500	
400	M59	ED-28 ED-37	Coated	Mog	11.3	20,000	36,000	22,500	
1000	M47	BT-36 BT-56	Clear	Mog	11.3	12,000	11,000 0	71,000	
35	S76	B-17	Clear	Medium		16,000	2,250	2,025	22
50	S68	B-17 or ED-17	Clear	Medium		24,000+	4,000	3,600	22
70	S62	B17 or ED-231/2	Clear	Mogul	73/4	24,000+	6,400	5,450	22

TABLE 26-13 Metal Halide and High-Pressure Sodium Hid Lamps[a] (*Continued*)

Watts	ANSI ballast type	Bulb	Bulb finish	Base	Maximum overall length, in	Vertical burning rated life, h	Initial	Mean	CRI
High Pressure Sodium Lamps—any burning position									
100	S54	B17 or ED-231/2	Clear	Mogul	73/4	24,000+	9,500	8,550	22
150	S55	B17 or ED-231/2	Clear	Mogul	73/4	24,000+	16,000	14,400	22
200	S66	ED-18	Clear	Mogul	93/4	24,000+	22,000	19,800	22
250	S50	ED-18	Clear	Mogul	93/4	24,000+	28,000	27,000	22
310	S67	ED-18	Clear	Mogul	93/4	24,000+	37,000	33,300	22
400	S51	ED-18	Clear	Mogul	93/4	24,000+	51,000	45,000	22
400	S51	ED-37	Diffuse	Mogul	11	24,000+	47,500	42,750	22
1000	S52	E-25	Clear	Mogul		24,000+	140,000	126,000	22

[a]Consult manufacturers' technical literature for current data, as values and types change frequently.
[c]Initial—after 100 h of burning; mean—for mercury and HPS at 50% of rated life, and for metal halide lamps, at 40% of rated life.
[d]Color rendering index.
[e]Check manufacturer's literature. May require enclosed luminaire.

TABLE 26-14 Advantages and Disadvantages of the Different Lamp Types

Light source	Advantages	Disadvantages
Incandescent general-service	Compact size, no ballast, good optical control, good color rendering, low first cost, dimmable, good lumen maintenance	Short life, low efficacy, radiant-heat effects
Tungsten-halogen incandescent	Compact, no ballast, good color rendering, moderate life, excellent optical control, dimmable, excellent lumen maintenance	Lamp handling is difficult during maintenance, high cost low efficacy, radiant-heat effects, operating temperature affects lamp life
Fluorescent	Linear shape, moderate cost, high efficacy long life, good lumen maintenance dimmable, rare-earth phosphers can give excellent color rendering	Ballast needed, optical control limited, ballasts may be noisy, ambient temperature affects light output and color
Improved-color mercury (HID) (Rarely used today)	Moderate efficacy, very long life, good lumen maintenance, burning position not critical, limited dimming range lamp and ballast	Starting takes 5–7 min, does not restart immediately, ballasts needed are large and may be noisy, relatively high cost of Variations in color, burn position very
Metal halide (HID)	High efficacy, good color rendering, medium life, good optional control, limited dimming range	important, does not restart immediately, ballasts large and may be noisy, high cost of lamp and ballast, starting takes 3–4 min
High-pressure sodium (HID)	Very high efficacy, long lamp life, excellent lumen maintenance, limited dimming range	Poor color rendering, does not restart immediately, ballasts large and may be noisy, high cost of lamp and ballast, starting takes 3–4 min, high luminance can cause control problems

REFERENCES ON HIGH-INTENSITY DISCHARGE LAMPS

1. ANSI, American National standard for Electric Lamps—High-Intensity Discharge Lamps, Method of Designation, C78.380-2005, American National Standards Institute, New York (also available at www.nema.org)
2. Nortrup, F., Kraska, Z., and Lou, S., Pulse Start of Metal Halide Lamps for Improved Lumen Maintenance, *J. Illum. Eng. Soc. N. Am.,* 1996, vol. 23, no. 2, pp. 113–116.

26.6 MISCELLANEOUS LAMPS

Low-Pressure Sodium Lamps. These are sodium vapor lamps in which the partial pressure of the vapor during operation does not exceed a few newtons per square meter. Their light output is almost monochromatic, consisting of a double line in the yellow region of the spectrum at 589 and 589.6 nm.

As with other electric discharge sources, a ballast is required. Starting time to full light output is 7 to 15 min, but the lamp will restart immediately after interruption of the power supply. The major application of these lamps is for area lighting and streetlighting where monochromatic yellow light is acceptable and a high luminous efficacy is required. Low-pressure sodium lamps are often applied in the vicinity of astronomical observatories, since the monochromatic light can be filtered out of telescope images.

Glow Lamps. When sufficient voltage is applied to electrodes sealed within a bulb containing neon, argon, or helium, light is produced at the negative electrode. On direct current, one cathode glows; on alternating current, the reversal is so rapid, both electrodes appear to glow. The range of glow lamps is 1/25 W to 3 W. Their useful life varies approximately as the inverse of the cube of the current. A glow lamp has a negative volt ampere characteristic; hence a limiting resistance is used in series with it. In conventional screw-base types, the resistor is concealed in the base. Average lamp life ranges between 7500 and 25,000 h.

Glow lamps have wide use in electronic circuitry, where their action is that of a practically instantaneous switch. At breakdown voltage, the lamp glows, and the switch is closed; at the extinguishing voltage, the lamp current drops to a fraction of its full value and may be considered as nonconducting, or open-circuit in certain circumstances. This on-off characteristic suits the glow lamp to the dichotomy of binary arithmetic as used in computers and logic circuitry in general. Other glow-lamp applications in electronic circuitry include oscillators, pulse generators, voltage regulators, and coupling networks.

Electroluminescent Lamps. This type of lamp is a thin-area source in which light is produced by a phosphor excited by a pulsating electric field. In essence, the lamp is a plate capacitor with a phosphor embedded in its dielectric and with one or both of its plates transparent. Green, blue, yellow, or white light may be produced by choice of phosphor. The green phosphor has the highest luminance. These lamps are available in ceramic and plastic form, are flexible or have a stiff backing, and are easily fabricated into simple or complex shapes. They have been used in decorative lighting, night lights, switchplates, instrument panels, clock faces, telephone dials, thermometers, and aircraft egress marking strips and signs. Their application is limited to locations where the general illumination is low.

Luminance varies with applied voltage, frequency, and temperature, as well as with the type of phosphor. Life is long and power consumption is low. There is no abrupt point at which the lamp fails; the time at which the luminance has fallen to 50% of initial is sometimes used as a measure of useful life. For the ceramic form, this is approximately 20,000 h at 120 V, 60 Hz. Approximate initial current and wattage values per square foot of lamp under these operating conditions are 60 mA and 3.5 W.

Black-Light Lamps.[1] Near-ultraviolet radiant energy (energy not visible to the human eye) causes certain materials to fluoresce or emit visible light. The normal human eye is sensitive only to radiant energy between 380 and 780 nm in wavelength. Thus, lamps that produce primarily near-ultraviolet radiant energy in the 320- and 380-nm range are popularly called *black* lights. This term is quite descriptive, since the ultraviolet energy from the light source cannot be seen by the human eye, but the effects of the radiation on special materials can be visually dramatic.

When black light is directed at a fluorescent material, an energy conversion takes place. The material or chemical sensitive to ultraviolet energy absorbs the energy, then reradiates it at longer wavelengths to which the eye is sensitive.

Mercury lamps with filters to absorb the visible light and transmit the near-ultraviolet are used for fluorescent effects. Called black-light lamps, they are generally enclosed in a red-purple filter glass bulb that looks black. Many materials fluoresce when irradiated by black-light lamps. They are used for theatrical and advertising effects, industrial and food inspection, detection of counterfeits and forgeries, medical diagnosis, insect traps, crime and vermin detection, laundry marking, and copying equipment.

Tubular sources designated as BLB lamps, such as 4-, 6-, and 8-W T-5, 15-W T-8, and 20- and 40-W T-12 lamps, have integral filters and may be operated with the same ballasts as corresponding fluorescent lamps.

The luminance of an irradiated fluorescent material is between 1 and 5 fL with printing inks and between 0.25 and 2.5 fL with interior paints, depending on the color. The apparent brightness increases considerably as the eyes become dark-adapted. Conversely, the effectiveness of black light is greatly reduced or entirely negated by a small amount of visible light.

Short-Arc Sources. Also called "compact arcs," these lamps have glass-enclosed arcs in bulbs containing mercury-argon, mercury-xenon, and xenon, to give the high brightness of carbon arcs, but without their dirty operational and maintenance problems. Principal applications are display systems, optical instruments, projectors, and searchlights.

Sulfur Lamps. Brilliant white light can be produced by using kitchen-grade microwave energy to excite a small quantity of sulfur within an argon-filled quartz sphere of golf-ball size.* The small sphere must be rotated at 300 to 600 r/min to keep the quartz from melting. The light output is said to be 125,000 to 175,000 lumens, and the target efficacy is 110 to 140 lumens per watt, with target input power of 800 to 1200 W.

Induction Lighting Systems. The discharge bulb contains mercury and an inert gas, and the inner surface is coated with phosphor, giving 3000 or 4000 K white light. An axially located power coupler serves as an antenna to radiate 2.65-MHz energy received via coaxial cable from an external high-frequency generator. Light output from the 85 system watts: initial 6000 lumens, mean 4800 lumens, CRI 80+, and rated life 100,000 h. (*Note:* this is the principle used for the 23-W R-25 bulb compact fluorescent lamp listed in Table 26-12.)

REFERENCES ON MISCELLANEOUS LAMPS

1. Kraehenbuehl, J. O., and Chanon, H. J., Technology of Brightness Production by Near-Ultraviolet Radiation, *Trans. IES*, Feb. 1941.
2. Cook, H., Stretching the Spectrum, *Building Design and Construction*, June 1998.

26.7 LUMINAIRES AND LIGHTING SYSTEMS

Luminaires. These are complete lighting units consisting of a lamp or lamps together with the parts designed to distribute the light, to position and protect the lamps, and to connect the lamps to the power supply. They are clarified in the CIE.

*Fusion Lighting Inc. Rockville, Md. See also Cook.[2]

(International Commission on Illumination) according to the percentage of light output above and below the horizontal as follows:

Direct	0%–10% upward, 90%–100% downward
Semidirect	10%–40% upward, 60%–09% downward
General diffuse	40%–60% upward, 40%–60% downward
Semi-indirect	60%–90% upward, 10%–40% downward
Indirect	90%–100% upward, 0%–10% downward

This classification system applies to all types of luminaires for general lighting in industrial, commercial, and residential applications.

Lighting systems are installations of one or more luminaires and are classified in a number of different ways. One of these is by CIE distribution type, since these different distributions provide different lighting quality, performance and space appearance.

Direct Lighting. When luminaires direct 90% to 100% of their output downward, they form a direct lighting system. The distribution may vary from widespread to highly concentrating, depending on the reflector material, finish, and contour, and on the shielding or control media employed. Troffers and downlights are two forms of direct luminaires.

Direct lighting units can have the highest utilization of all types, but this utilization may be reduced in varying degrees by brightness-control media required to minimize direct glare. Veiling reflections and shadows may be excessive unless the distribution and location of luminaires are designed to reduce these effects. Large-area units are generally also advantageous since they soften shadows.

Luminous ceilings, louvered ceilings, and large-area modular lighting elements are forms of direct lighting having characteristics similar to those of indirect lighting discussed in paragraphs below. Luminous ceilings may be difficult to apply at low power densities.

Semidirect Lighting. The distribution from semidirect units is predominantly downward (60% to 90%) but with a small upward component to illuminate the ceiling and upper walls. The characteristics are essentially the same as for direct lighting except that the upward component will tend to soften shadows and improve room brightness. Care should be exercised with close-to ceiling mounting of some types of prevent overly bright ceilings directly above the luminaire. Utilization can approach, or even sometimes exceed, that of well-shielded direct units.

General Diffuse Lighting. When downward and upward components of light from surface-mounted and suspended luminaires are about equal (each 40% to 60% of total luminaire output), the system is classified as general diffuse. General-diffuse units combine the characteristics of direct lighting described above and those of indirect lighting described below. Utilization is somewhat lower than for direct or semidirect units, but it is still quite good in rooms with high-reflectance surfaces. Brightness relationships throughout the room are generally good, and shadows form the direct component are softened by the upward light reflected from the ceiling.

Direct-indirect is a special (non-CIE) category within this classification for luminaires which emit very little light at angles near the horizontal. Since this characteristic results in lower luminances in the direct-glare zone, direct-indirect luminaires are often more suitable than general-diffuse luminaires which distribute the light about equally in all directions, especially in spaces involving critical and prolonged seeing.

Semi-Indirect Lighting. Lighting systems which emit 60% to 90% of their output upward are defined as semi-indirect. The characteristics of semi-indirect lighting are similar to those of indirect systems discussed below except that the downward component usually produces a luminaire luminance that closely matches that of the ceiling. However, if the downward component becomes too great and is not properly controlled, direct or reflected glare may result.

Indirect Lighting. Lighting systems classified as indirect are those which direct 90% to 100% of the light upward to the ceiling and upper sidewalls. In a well-designed installation, the entire ceiling becomes the primary source of illumination, and shadows are virtually eliminated. Also, since the luminaires direct very little light downward, both direct and reflected glare will be minimized if the installation is well planned. Luminaires whose luminance approximates that of the ceiling have some advantages in this respect. It is also important to suspend the luminaires a sufficient distance below the ceiling to obtain reasonable uniformity of ceiling luminance without excessive luminance immediately above the liminaires.

Since with indirect lighting the ceiling and upper walls must reflect light to the work plane, it is essential that these surfaces have high diffuse reflectances, but low specular reflectances. Care is needed to prevent overall ceiling luminance from becoming too high and thus glaring.

Control of Light Distribution. This is usually accomplished through reflection, refraction, transmission, absorption, and diffusion using glasses, plastics, metals, and woods, of various shapes, reflectance, transmittance, absorptance, polarization, and finish. Reflector contour shapes include parabolic, ellipsoidal, hyperbolic, and spherical. Refractors (lenses) utilizing prisms, cones, and spherical shapes are commonly used to produce a wide range of light-controlling devices. Flat or contoured diffusers are used to diffuse, color, or polarize the light according to the lighting needs.

Lighting systems are also classified in accordance with their layout or location with respect to the visual task or object lighted—general lighting, localized general lighting, and local (supplementary) lighting.

General Lighting. Lighting systems that provide an approximately uniform level of illumination on the work plane over the entire area are called general lighting systems. The luminaires are usually arranged in a symmetrical plan fitted into the physical characteristics of the area and blend well with the room architecture. They are relatively simple to install and require no coordination with furniture or machinery that may not be in place at the time of the installation. Perhaps the greatest advantage of general lighting systems is that they permit complete flexibility in task location. Since they illuminate the entire space to the same level, they may provide more light than is necessary in certain parts of the room, and therefore consume more energy than needed.

Localized General Lighting. A localized general lighting system consists of a functional arrangement of luminaires with respect to the visual task or work areas. It also provides illumination for the entire room area. Such a lighting system requires special coordination in installation and careful consideration to ensure adequate general lighting for the room. This system has the advantages of better utilization of the light on the work area and the opportunity to locate the luminaires so that annoying shadows and direct and reflected glare are prevented.

Local Lighting. A local lighting system provides lighting only over a relatively small area occupied by the task and its immediate surround. The illumination may be from luminaires mounted near the task (task lighting) or from remote spotlights. It is an economical means of providing higher illumination levels over a small area, and it usually permits some adjustment of the lighting to suit the requirements of the individual. Improper adjustments may, however, cause annoying glare for nearby workers. Local lighting, by itself, is seldom desirable. To prevent excessive changes in adaptation, it should be used in conjunction with general lighting that is at least 20% of the local lighting level; it then becomes *supplementary lighting*. This combination of a local and general system is generally referred to as task-ambient lighting. Task ambient lighting can provide excellent energy efficiency because high illuminance levels are applied only where needed.

Application Considerations. There are many utilitarian, esthetic, energy-efficiency, and economic considerations that influence the selection of light sources, luminaires, and lighting system for interior spaces. The choices are simple for illuminating a janitorial closet, but increasingly difficult for areas requiring reasonably inconspicuous luminaires which provide the quantity and quality of light that enable rapid and accurate seeing of critical and prolonged visual tasks.

Changes in building design and escalating construction costs have combined to lower ceiling heights in typical commercial and institutional spaces. The expansion of computers and word processors containing visual display terminals (VDTs) brought many shiny, curved, near-vertical tasks of relatively low contrast into the visual environment. The parabolic troffer was for many years the lighting solution in office environments. The general environment provided by these luminaires often appears somewhat dark due to the low luminance of the ceiling plane. Appropriately configured pendant-mounted lighting systems with some amount of uplight generally provide higher quality office lighting, since these systems can both limit veiling reflections and provide higher overall space brightness.

Commercial and Institutional Interiors. In such spaces, the need exists to limit direct glare, reflected glare, and veiling reflections by selecting luminaires with appropriate luminous intensity distributions and finishes of furniture and room surfaces, as well as task reflectances, that provide appropriate illuminance levels and achieve luminance ratios within recommended limits (no greater than 3:1 or 1:3 for the task to the near surround and no more than 10:1 or 1:10 between the task and the far surround.)

As discussed above, a widely used luminaire type is the recessed troffer, with prismatic, parabola-shaped or other louvered enclosures. Standard sizes are 1 × 4 ft, 20 × 48 in, 2 × 2 ft, and 2 × 4 ft. The 3- and 4-in deep parabolic louvers are popular when critical and prolonged seeing tasks are involved, or where inconspicuous lighting systems are desired.

Pendant luminaires are available in a variety of distributions, shapes, and sizes. T8 lamps are the most common lamps in these luminaires, but the small diameter of T5 lamps generally provides for better optical control and permits smaller pendant luminaire cross sections.

Hard-metric* recessed luminaires are available as a result of the 1994 policy requiring them in future construction projects of all U.S. federal agencies. Later the 1996 Savings in Construction Act passed in the 105th Congress mandated that hard-metric luminaires must not be more costly than soft-metric[†] models.

Computer monitors are increasingly available with finishes that are less susceptible to reflected glare, and as a result, luminaires such as recessed direct-indirect troffers are being promoted by some manufacturers to provide environments with more vertical illuminance for walls and other surfaces. These large square of rectangular units, with all the components above the ceiling plane, contain shielded fluorescent-lamp luminaire elements positioned to direct all or most of the light to white-finished curved or sloping sides, but are actually a CIE direct luminaire distribution. The tendency of these recessed direct-indirect luminaires to produce direct or reflected glare is similar to that of lensed luminaires, and is a function of their luminance.

Specialized models of small- and large-cell parabolic louvers in recessed and surface-mounted luminaires are designed to meet standards of the Illuminating Engineering Society of North America for spaces with VDTs. Parabolic luminaires expected to be used with T-8, high-lumen compact, and other rare-earth phosphor fluorescent lamps should be equipped with low-iridescence aluminum louvers.

The National Electrical Manufacturers Association provides dimensions to help ensure compatibility with conventional generic systems for ceiling suspension. Dimensions are provided for a variety of different ceiling systems, along with cross-section details showing how the luminaires interface with the ceiling systems[1].

Combining Lighting and Air Conditioning. Families of recessed fluorescent-lamp luminaires (troffers) are designed to provide the additional function of bringing cool or warm air into interior spaces and to take air out of those spaces, thus eliminating or minimizing the need for supply air diffusers and return air grilles. The air-supply function is accomplished by slots along the bottom edge of the luminaire sides, above which sheet-metal air connectors (frequently called *airboots*—see Fig. 26-23*a*) spread out the conditioned air received by 5-, 6-, and 7-in flexible ducts form larger plenum ducts. Air connectors can be above one or both side slots, can be internally insulated or uninsulated,

*Hard-metric: even multiples of 100 mm (e.g., 600 × 1200 mm).

[†]Soft-metric: a simple conversion of inches to metric equivalents [e.g., 609.6 × 1219.2 mm (24 × 48 in)].

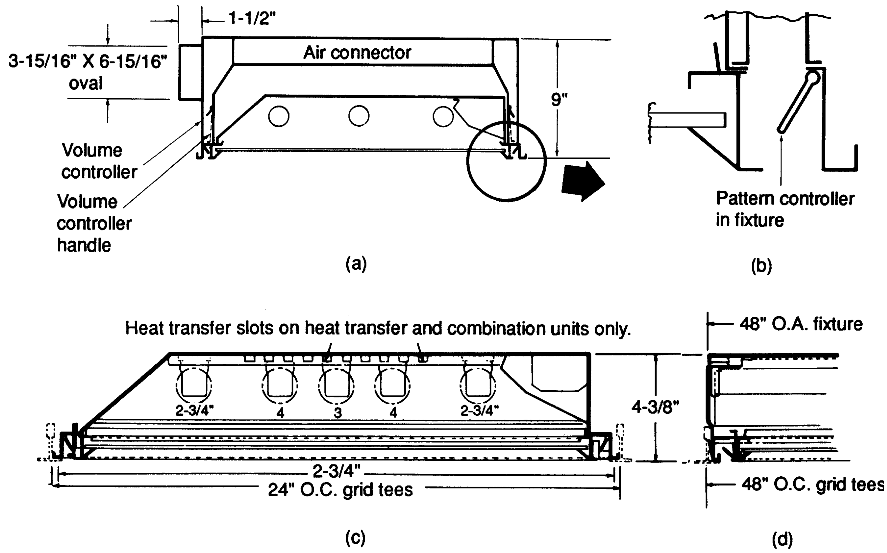


FIGURE 26-23 Typical cross sections and dimensions of (*a*) air connectors (boots), (*b*) air-supply pattern control blade, (*c*) cross, and (*d*) end views of air-handling supply and/or heat-transfer-recessed fluorescent luminaire.

and can have volume-control dampers. The luminaire side slots can include air pattern-control blades (Fig. 26-23*b*), which are field positioned to be closed (static), at 45° for horizontal air supply, and vertical (fully open) for vertical air supply. Side slots may also be used to return air to the plenum, or to connect directly to a ducted air return system.

Room air can also be pulled through the lamp compartment of troffers by creating negative air pressure in the plenum and building in suitable air entry and air exist openings in the luminaire. By this process, considerable lamp and ballast heat is then expelled through top openings into the plenum. Called *heat transfer* or *heat exhaust*, such luminaires can significantly reduce the heat sent into the occupied space below, increase light output because of the more optimum temperature inside the lamp compartment, and provide a thermal environment for more favorable ballast life.

Air-handling luminaires can combine both the air supply and heat-transfer return air functions, called *combination units* in Fig. 26-23*c*, which shows the top heat-transfer openings, and the two end openings (Fig. 26-23*d*), which allow room air to enter the lamp compartment.

Iuminaire mounting issues. Incandescent- and HID-lamp luminaires for surface and recessed applications are widely used for downlighting, wall washing, etc. Surface or recessed track with suspended, adjustable incandescent luminaires is popular for display and localized lighting in stores, restaurants, art galleries, museums, etc. See the latest edition of the NEC for special requirements, such as thermal protection for incandescent and HID recessed luminaires, minimum clearances for clothes closets, and wattage and temperature markings.

Fluorescent-lamp luminaires designed for surface mounting are of the prismatic wraparound, metal-box, or bare-lamp (strip) types. If such luminaires are to be installed on combustible low-density cellulose fiberboard, they must be listed for that condition, or be spaced at least $1^1/_2$ in from the surface of the fiberboard.

Industrial Interiors. Luminaires with fluorescent lamps are often preferred for low-bay industrial spaces, generally involving mounting heights 12 ft or less above the floor. Eight-foot units with center-V reflectors and 96 in 800 mA "high output" lamps frequently produce lower overall cost of light than slimline and 1500-mA models, although the extra-high-output type normally results in

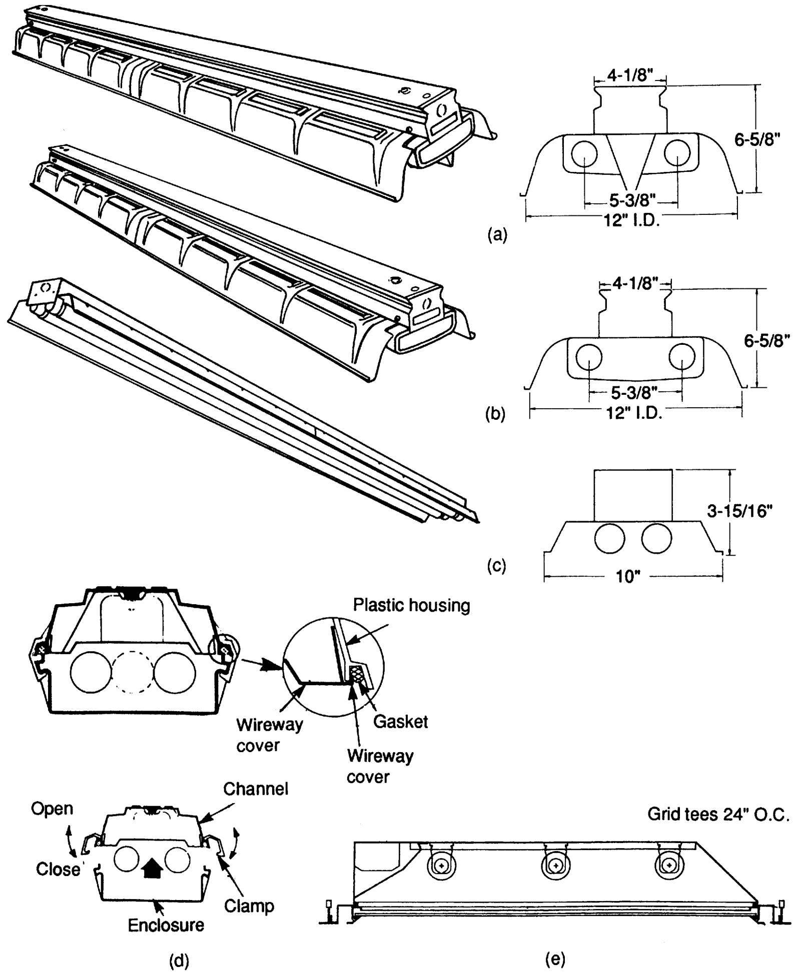


FIGURE 26-24 Industrial fluorescent-lamp luminaire types for nonhazardous areas.

most favorable initial cost. Units with four 4-ft lamps are popular with many plant engineers and maintenance supervisors because of the shorter bulb length. Figure 26-24 shows the wide top slots that give approximately 25% upward light, and the center-V providing greater reflector rigidity and 30° crosswise shielding, which improves visual comfort by reducing direct glare, especially when luminaires are oriented so the long dimension is at right angles to the predominant direction of the worker's line of sight (called *crosswise viewing*). Other models without center-V reflectors frequently have narrower top slots giving approximately 10% upward light and 13° crosswise shielding (Fig. 26-24*b*). Economical luminaires have more shallow 10% uplight or closed-top reflectors (Fig. 26-24*c*). Steel reflectors are most common, with porcelain-or baked-enamel finish.

T5 fluorescent lamps are also applied in industrial lighting where high output linear lamps can be grouped together in a single luminaire to provide sufficient light output for use in high bay applications where HID luminaires are typically applied. Traditional high-bay luminaires with round aluminum, glass, or acrylic reflectors (Fig. 26-25) are also available with multiple compact fluorescent lamps instead of HID lamps. These systems can provide multiple light levels through switching of

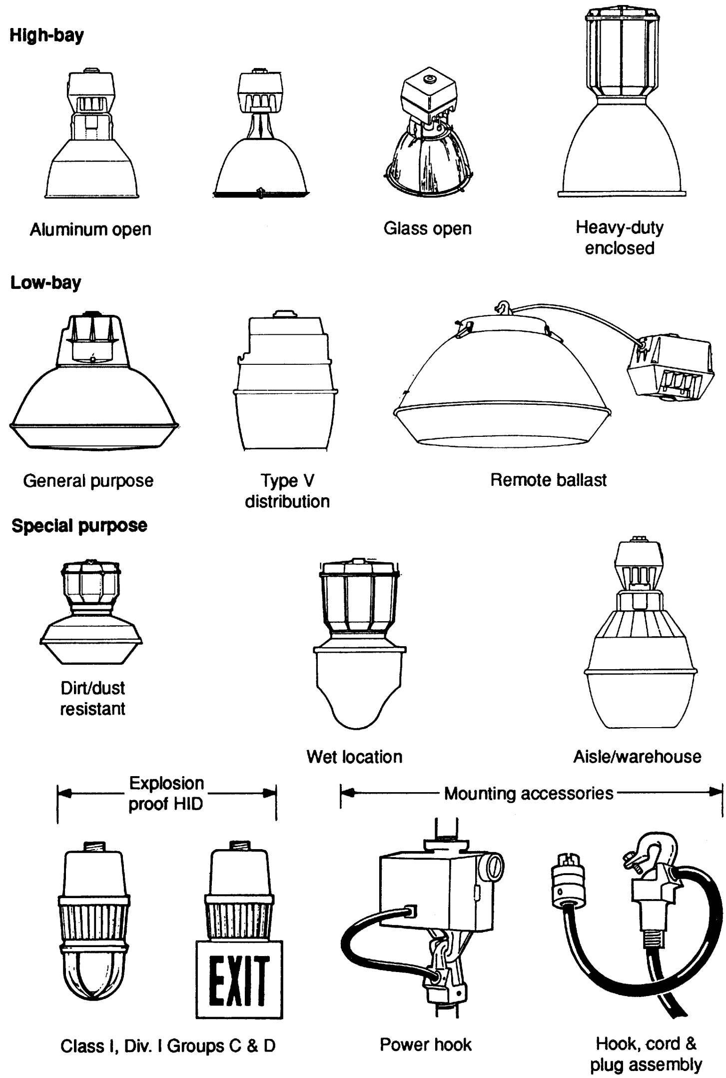


FIGURE 26-25 Industrial HID and incandescent lamp luminaire types and mounting accessories, and two explosion-proof models.

different groups of lamps in the same luminaire, and have been applied in warehouses, stores, gymnasiums and numerous other applications.

Specialized nonhazardous area industrial fluorescent luminaires include enclosed types for dirt- and moisture-resistant applications (Fig. 26-24*d*), recessed troffers for clean rooms (Fig. 26-24*e*), and compact fluorescent supplementary lighting units. Classifications of hazardous areas are defined in the NEC. Advice of the insurance carriers for the hazardous spaces involved is recommended to confirm appropriate classification before installing fluorescent, HID, or incandescent luminaires designed for hazardous areas.

Metal halide and high-pressure sodium industrial luminaires are prevalent for production, warehouse, and other industrial areas. Metal-halide lamps have color characteristics that are superior to high-pressure sodium. High-pressure sodium systems generally result in more rapid payback for situations where minimum energy consumption is the goal and color rendition is an insignificant factor. However, the warm color of high-pressure sodium may be judged undesirable, especially if the visual tasks are such that the high-pressure sodium light would adversely affect seeing speed or accuracy. For example, workers may be less able to see small defects in brass and copper tasks, compared with other light sources.

Figure 26-25 shows typical HID industrial luminaires. High-bay units are generally for mounting heights of 18 ft or more above the floor, where cranes or maintenance catwalks simplify the task of relamping and reflector cleaning. Low-bay models are often selected when luminaires need to be positioned 13 to 18 ft above the floor, and periodic maintenance can be performed with portable scaffolding or lift trucks. Both the high- and low-bay types can employ remote ballasts. Integral or remote, the ballast must be suitable for the ambient temperature involved at the luminaire location.

A variety of HID luminaires are designed for specific purpose, and manufacturers provide mounting and wiring accessories that expedite installation and maintenance. Fuses (single for 120- and 277-V circuits, and double for 208, 240, and 480 V) can be located either in the ballast assembly or inside the power hook. Since regular HID lamps will not immediately restart after a momentary power interruption or a severe voltage dip, a tungsten halogen lamp inside the reflector can provide standby illumination for safety during the few minutes the high-pressure sodium or metal halide lamps are out. A "hot restrike" option is available for metal halide and high-pressure sodium systems. All HID ballasts produce some noise—for quieter operation of HID luminaires, encased and potted ballasts are available.

Merchandise and Display Lighting. Systems achieving the "3 A's" of store lighting (attraction, appraisal, and atmosphere) utilize many luminaire types to (1) call attention to featured items; (2) reveal their inherent characteristics and color; and (3) provide an appropriate, stimulating environment. Some stores may have free-standing racks and open cases, where essentially uniform illuminance is suitable. Or important sales areas may be arranged so that luminaires could be selected or grouped to produce nonuniform illuminance. Still others may have showcases with glass tops and fronts, where certain types or positions of overhead luminaires might cause bothersome reflections that reduce the visibility of merchandise inside the case. Niches can have built-in linear or incandescent lighting, and track, surface, and recessed luminaires are widely employed for vertical displays.

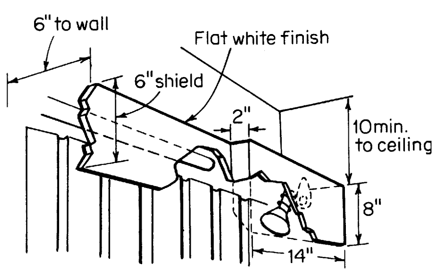


FIGURE 26-26 Valance and spotlight.

Figure 26-26 illustrates a wall-lighting valance that has a spotlight in the end section, which is separately controlled for highlighting. The shelf valance (Fig. 26-27) is useful for books, merchandise, and art objects, and for tasks not affected by veiling reflections. The room divider (Fig. 26-28) is applicable where unobtrusive divisions of space are desirable. Consult other references[2] for additional solutions.

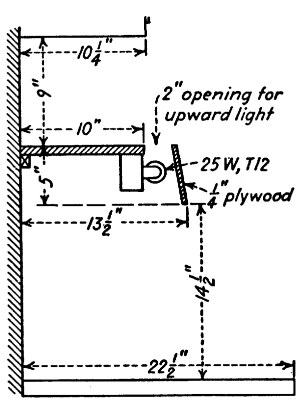


FIGURE 26-27 Shelf valance.

Structural Lighting. Lighting systems that form a substantial part of the structure of a building, as distinguished from individual or groups of luminaires suspended from the ceiling or bracketed to walls, are examples of structural lighting.

A luminous ceiling of suspended plastic or other material is one such form (Fig. 26-29). Luminous ceilings are used less frequently today due to their relative inefficiency. One of the more common forms of structural lighting used today is an illuminated ceiling cove, where a recessed ceiling area is indirectly illuminated from a small shelf around the perimeter through the use of either bare lamp strips, or reflector luminaires designed to direct light out onto the ceiling. In most cove applications, the luminaires or lamps are hidden by the lip of the cove.

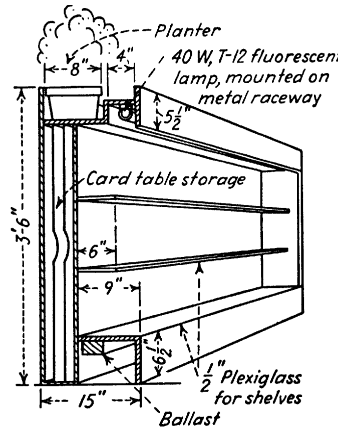


FIGURE 26-28 Room divider.

Codes and Standards. Local codes, national codes, international codes, federal standards, professional standards, and manufacturers' *standards* relate to specific requirements, which must be met in the construction and installation of a luminaire. Standards usually relate to minimum requirements (safety, construction, or performance), which may be exceeded to provide a better product.

Some codes and standards deal with fire and safety (electrical, mechanical, and thermal); others relate to performance (photometric) and construction (materials and finishes). They will vary to some extent depending on geographic location and end use of equipment. Conformance to the appropriate set of specifications is often determined by certified laboratory tests. Certification is often denoted by an identifying label. Local inspection agencies may or may not rely on conformance to national, federal, or industrial codes and standards.

Information regarding local codes may be obtained from electrical inspection departments. Several other code jurisdictions may apply to regional and state codes. The National Electrical Code (NEC) and similar codes in most major countries throughout the world state specific electrical requirements, which must be met by all electrical equipment, including luminaires. They have been developed by safety protection and inspection agencies in conjunction with fire-protection agencies.

The Underwriters' Laboratories, Inc. (UL) and similar groups in other countries publish minimum safety standards for electrical and associated products which are in conformance with the respective electrical codes of their country. They have testing laboratories to which equipment must be submitted for listing. Most manufacturers design luminaires to meet these standards.

The Life Safety Code[3] (National Fire Protection Association 101) covers the requirements of exit signs and emergency lighting. Many models of standard and emergency exit signs are designed to meet NFPA-101 luminance standards, and retrofit kits permit easy conversion to more efficient sources. Signs equipped with battery packs must remain lighted a minimum of 90 min when their normal electric power fails. Engineers are encouraged to check for enactments of specific states regarding the required positions (such as requirements for not only the overdoor location, but also a low-level exit sign) and performance attributes (letter configurations, color, etc.) of exit signs.

Building codes also have requirements for egress lighting. The NFPA-101 *Life Safety Code* identifies occupancies that require emergency lighting. Most codes call for lighting the exit passageways

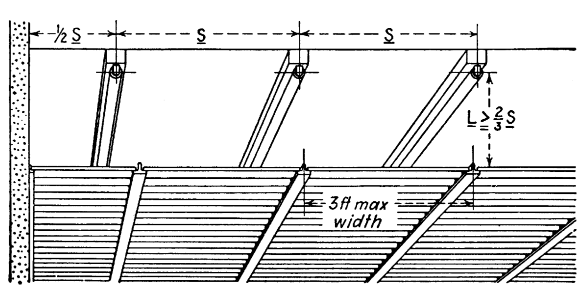


FIGURE 26-29 Luminous ceiling.

to a minimum illuminance of 1 fc on the floor for a period of $1^1/_2$ h after normal power fails. Luminaires containing fluorescent emergency ballasts having suitable, continuously charged battery packs can be positioned to meet requirements with acceptable uniformity. Ballasts providing instant restrike of high-pressure sodium luminaires are also available.

In addition to construction and safety standards, there are national and state energy-efficient lighting regulations which followed the federal Energy Policy Act of 1992. For commercial buildings, provisions of the latest ASHRAE/IESNA Standard 90.1, 2004 may prevail. Except for emergency lighting powered by battery, generator, or other alternate power source, which is automatically off during normal building operation, under this standard the lighting system must adhere to the following.

1. In buildings larger than 465 m^2 (5000 ft^2), automatic lighting shutoff of building lighting is required in all spaces, except for lighting specifically designed for 24-h use. This may be accomplished on a scheduled shutoff basis or through occupant sensors.

An independent schedule must be provided for areas up to 2323 m^2 (25,000 ft^2), but not more than one floor.

Spaces with ceiling-height partitions must have at least one control device to independently control the general lighting within the space, actuated manually by an occupant or automatically by sensing an occupant.

2. Luminaires for one or three fluorescent lamps greater than 30 W each must use two-lamp tandem-wired ballasts instead of single-lamp ballasts, when two or more luminaires are in the same space, and on the same control device. This does not apply to recessed luminaires more than 3 m (10 ft) apart measured center to center, individually mounted surface or pendant luminaires, luminaires with single-lamp electronic ballasts, luminaires with no available pair, or luminaires on emergency circuits.

3. Internally illuminating exit signs shall not exceed 5 W per face.

The 90.1 Standard includes important sections on installed interior lighting power and luminaire wattage. Power limits are computed by either the "building are" or the space-by-space method. Tables give maximum lighting power density (LPD) in watts per square foot for typical interior spaces. For example, enclosed offices are 1.1 W/ft^2, open-plan offices are 1.1 W/ft^2, classrooms/lecture/training spaces are 1.4 W/ft^2, and restrooms are 0.9 W/ft^2, and so on. The standard should be reviewed so that various exceptions can be considered. Other examples are for exterior lighting of a building entrance with canopy, 1.25 W/ft^2; without canopy, 30 W/ft of door width; and 0.2 W/ft^2 for lighting vertical facade area, or 5 W/ft of illuminated wall or surface length.

Legislation requiring adherence to ASHRAE/IESNA 90.1 varies considerably among the states. Engineers are urged to check with appropriate agencies for latest information. Certain states mandate statewide conformance to state-developed energy codes for commercial buildings which meet or exceed ASHRAE/IES 90.1.

The desire of commercial, industrial, and institutional facility managers to reduce costs while maintaining and/or improving the visual environment for occupants also adds to the priority placed on energy conservation in lighting systems. Fortunately, major progress has been made in rating the combined efficacy of luminaires and the lamps and ballasts they contain. Also, modest-cost lighting control devices are available, which quickly pay for themselves by saving energy. See further information, see the subsection on energy conservation later in Sec. 26.9. There are also opportunities and benefits from integrating lighting with other building systems.

REFERENCES ON LUMINAIRES AND LIGHTING SYSTEMS

1. NEMA LE 4-2001, Recessed Luminaires; Ceiling Compatibility, National Electrical Manufacturers Association, www.nema.org
2. RP-2-015, *Lighting Merchandising Areas (A Store Lighting Guide)*, Illuminating Engineering Society of North America.

3. NFPA 101, *Life Safety Code*, 2006 ed., National Fire Protection Association.

4. ASHRAE/IESNA Standard 90.1/2004, *Energy Standard for Buildings Except Low-Rise Residential Building*, American Society of Heating, Refrigerating and Air-Conditioning Engineers & Illuminating Engineering Society of North America, 2004.

26.8 LUMINAIRE PHOTOMETRIC DATA

Types of Photometric Data. Luminaire manufacturers provide laboratory-measured and computed photometric and physical data required for lighting system analysis. The physical data include dimensions, finishes, light-control materials, number of lamps, types of lamps, etc. The photometric data include a luminous intensity (candela) distribution, zonal lumens, luminaire efficiency, coefficients of utilization, luminances, and for some systems, visual comfort probability (VCP) values.

Luminous Intensity (Candela) Distribution. A luminous intensity or candela distribution curve is a graphical presentation of the distribution of light about a lamp or luminaire. Such presentations contribute valuable information to guide the lighting designer in determining the suitability of lighting equipment for application in various fields. As a background for using distribution curves, it is first necessary to understand how they are obtained.

The candelas in any direction from a light source equals the illumination produced on a plane at right angles to the light rays times the square of the distance from the lamp to the point of measurement ($I = E \times D^2$). For accurate measurements the distance should be at least 5 times the largest dimension of the source.

To get sufficient data for a lamp or luminaire, readings are usually taken at a series of horizontal and vertical angles around a luminaire. Vertical angles may range from 0° (straight down) through 180° (straight up) with a spacing of 2.5°, 5° or 10°, while for most indoor luminaires, readings generally have a horizontal angle spacing of 22.5°. Horizontal symmetry and lack of uplight or downlight may limit the number of angles that are needed to fully describe a luminaire's intensity distribution. The luminous intensity values are typically plotted on a polar plot (Fig. 26-30) and may be shown for more than one horizontal angle when the distribution varies across the different axes. The line connecting a series of such points forms the intensity distribution curve. For concentrated light sources such as searchlights and spotlights, photometric readings are often required to be 1° or 2° apart.

The luminous intensity values reported for a particular manufacturer's luminaire apply to a particular lamp and its rated lumen output. If a lamp of identical physical size and light distribution, but having different rated lumens is applied in the luminaire, the luminous intensity values can simply be modified by the ratio of the rated lamp lumens for the new lamp

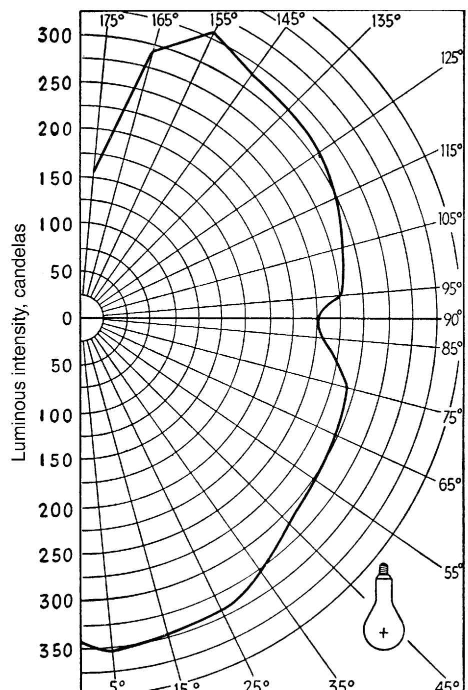


FIGURE 26-30 Distribution curve of a 3700-lm 200-W incandescent lamp. For convenience, the curves of axially symmetrical-type units show a single angle that is repeated at all other horizontal angle (0° or 360°). Luminous intensity in candelas is often referred to as a *candlepower distribution curve*.

versus the photometric test lamp. If the lamp is of different size or different distribution, then a new photometric report is needed to accurately quantify the luminaire's modified intensity distribution.

Zonal and Total Lumens. The total lumens of a light source are obtained by adding the lumens in the various angular zones. The zonal lumen values are readily computed by multiplying the average candelas around a lamp or luminaire at the center of a vertical angle zone by the zonal constant for that zone. For example, in Table 26-15, at 55°, which is the center of the 50° to 60° zone, the 300-cd reading from Fig. 26-37 is multiplied by 0.90, the zonal constant for this zone, giving 270 lm in that zone for the 200-W lamp. Zonal constants for the usual 18 test zones are simply numerical values of the solid angle subtended by that zone, which is the area in square feet within each zone on the surface of a sphere having a radius of 1 ft.

Luminaire Efficiency. The efficiency of a luminaire is expressed in terms of its lumen output divided by the rated lumens of the lamp or lamps. For example, if the 200-W lamp in Fig. 26-30 is placed in a white-glass enclosing globe whose candlepower distribution shows an output of 2960 lm, the efficiency of the luminaire is 2960 divided by 3700, or 80%.

Coefficients of Utilization. A coefficient of utilization (CU) is a number used in general lighting calculations and represents the ratio of the lumens received on the work plane from a uniform layout of luminaires to the lumens emitted by the luminaires' lamps alone. CUs are calculated using a uniform method and are tabulated for various room shapes and reflectances and are applied in the lumen method (discussed later).

Luminance and VCP. Average luminances calculated at various luminaire angles are part of direct-glare evaluations of VCP. Calculated VCP ratings of lighting installations using specific luminaire photometry, for various size rooms of given reflectances and illumination levels, may be reported. These ratings express the percentage of people who, if seated in the most undesirable location

TABLE 26-15 Tabulation of Candelas and Lumens in the Various Zones in the Curve of Fig. 26-37 (Computation of lumens from candelas)

Angle	Candelas	Zone	Zonal constant	Zonal lumens
0	347			
5	352	0–10	0.10	35
15	348	10–20	0.28	97
25	342	20–30	0.46	157
35	326	30–40	0.63	205
45	307	40–50	0.77	236
55	300	50–60	0.90	270
65	288	60–70	0.99	285
75	285	70–80	1.06	302
85	259	80–90	1.09	282
90	257			
95	271	90–100	1.09	295
105	278	100–110	1.06	295
115	290	110–120	0.99	287
125	307	120–130	0.90	276
135	308	130–140	0.77	237
145	313	140–150	0.63	197
155	329	150–160	0.46	151
165	280	160–170	0.28	78
175	153	170–180	0.10	15
Total lumens	…	…	…	3700

within the installation (4 ft from the center of the back wall and 4 ft above the floor), will be expected to find it acceptable. VCP was developed to quantify the glare of lensed fluorescent troffers, although it is often also applied to parabolic troffers. It cannot be applied to small downlights, HID luminaires, or luminaires with uplight, so its utility is somewhat limited. The CIE provides an alternate glare evaluation system, unified glare rating (UGR), that may apply to a wider range of luminaire types, but it is not widely used in the United States.

Luminaire Spacing Criteria (SC). The maximum permissible spacing for a luminaire is given in the Photometric Report provided by the manufacturer. These spacing limitations are related to the mounting height (usually above the work plane) and apply only to direct and semidirect luminaires. The maximum luminaire spacing center-to-center should not exceed the SC value times the mounting height (MH) and are often listed for the directions along and across the lamps for linear fluorescent luminaires. Observance of such limitations usually will ensure satisfactory uniformity of illumination throughout the major portion of the room so that all parts of the area will be equally suitable for the intended use. Peripheral areas may require special treatment, as indicated below. Illuminance is usually considered uniform if the maximum and minimum values are within plus or minus one-sixth of the average illuminance in the area. Closer spacing than indicated by the spacing criterion improves uniformity and reduces shadows. For luminaires with a large portion of light directed upward, computer software that performs point-by-point calculations can be applied to assess system uniformity.

The distance between luminaires and the wall should not exceed one-half the distance between luminaires. Where desks or benches might be located along the wall, the distance between luminaires and the wall should not exceed $2^1/_2$ ft. Likewise, the ends of continuous rows of fluorescent luminaires should preferably be within 6 to 12 in of the wall, unless this would provide a high luminance "scallop" on the adjacent wall. Additional luminaires or luminaires having a greater number of lamps may be required adjacent to the walls, particularly where walls have low reflectance. Where direct and semidirect luminaires are used under such conditions, the perimeter luminaires should be carefully located to avoid shadows on the task from the worker.

26.9 LIGHTING DESIGN

General Process. The lighting-design process can be broken down into four major parts: (1) determining the project goals, (2) collecting the space criteria and selecting the desired lighting system performance criteria, (3) making design decisions, and (4) evaluating the completed project. The design process may require changes in decisions (3) until the project goals have been satisfied.

Project Goals. These are the conceptual objectives of a whole project. The designer must have knowledge of the appearance requirements of the job, including the style of the architecture; whether a corporate image is to be projected; whether construction is for long life or of a speculative nature; whether young or old persons will be using the facility; and any other objectives of the project which will affect the lighting design.

Space Criteria. These are criteria which originate from outside the engineering discipline. They include identification of visual tasks and their locations and frequency, space dimensions, workplane locations, reflectances, daylight apertures, space temperature and dirt conditions, voltage, operating schedules, cost budgets, power and energy budgets, and codes. These criteria should represent the best available information which can be obtained from the owner, architect, or designer.

Performance Criteria. These are criteria which originate from within the illuminating engineering discipline. They include illuminance levels for task performance and safety, visual comfort data, luminance ratios, other quality considerations. These criteria can be found in various American

TABLE 26-16 IESNA Illuminance Categories and their recommended values.

Illuminance category	Footcandles	Lux
	Occasional tasks	
A	3	30
B	5	50
C	10	100
	Common tasks	
D	30	300
E	50	500
F	100	1,000
	Special tasks	
G	300–1,000	3,000–10,000

Note: The letter categories have the following descriptors:
A Public spaces with dark surroundings
B Simple orientation for short, temporary visits
C Working spaces where visual tasks are only occasionally performed
D Performance of visual tasks of high contrast or large size
E Performance of visual tasks of medium contrast or small size
F Performance of visual tasks of low contrast or very small size
G Performance of visual tasks near threshold

National Standards Institute practices (Office,[1] School,[2] Industrial,[3] Roadway[4]), in IES recommended practices and committee reports and in the IES Lighting Handbook.

Lighting design has long been known to involve more than selecting and positioning luminaires to provide recommended illuminances on visual tasks, wherever they are located in horizontal, inclined, or vertical work planes. The Committee on Quality of the Visual Environment of the Illuminating Engineering Society of North America, and that society's application committees, have developed matrices of quality issues for various types of interiors (Table 26-16 and Fig. 26-31). These matrices remind engineers and designers of specific aspects (source/task/eye geometry, color appearance, flicker, direct glare, etc.[1]) which merit very important, important, somewhat important, or no consideration for the areas involved. For example. industrial spaces might have matrices with sections for basic tasks identified as raw-materials processing, materials handling, component manufacture, machining, assembly, warehouse and storage, inspection, service spaces, welding, and manual crafting. Some applications could be subdivided—for instance, the warehouse and storage section might have three rows of criteria ratings, indicating the priority of consideration for inactive, active and active (small items). The assembly section could have rows for simple, moderately difficult, difficult, very difficult, and exacting assembly. These matrices appear in the IES Lighting Handbook.

Design Decisions. All the above criteria along with the project goals need to be evaluated when making major design decisions. These major decisions include light-source selection, luminaire selection and mounting, maintenance-procedure determination, calculation-methods selection, luminaire-layout arrangement, and control planning (switching, dimming). Decisions involving tradeoffs and compromises are often necessary.

Evaluation. No design process is complete until one has evaluated the results to see that the project goals have been met. Computer software and hand calculation methods exist to provide detailed point-by-point work plane or room surface analysis grids or average work plane illuminance for evaluating lighting layouts during the design phase. Many of these programs are also able to provide

I. Interior locations and tasks	Very important / Important / Somewhat important / Blank = Not important or not applicable																						
Design issues	Appearance of space and luminaires	Color appearance (and color contrast)	Daylighting integration and control	Direct glare	Flicker (and strobe)	Light distribution on surfaces	Light distribution on task plane (Uniformity)	Luminances of room surfaces	Modeling of faces or objects	Point(s) of interest	Reflected glare	Shadows	Source/task/eye geometry	Sparkle/desirable reflected highlights	Surface characteristics	System control and flexibility	Special considerations	Notes on special considerations	Illuminance (horizontal)	Category or value (lux)	Illuminance (vertical)	Category or value (lux)	Notes on Illuminance - see end of section	Reference chapter(s)
Offices (13)																								Ch. 11
Filing (see reading)	Somewhat important	Somewhat important		Important	Somewhat important	Important	Important	Somewhat important	Somewhat important		Somewhat important	Somewhat important	Somewhat important		Somewhat important				Important	E	Very important	C		
General and private offices (see reading)																								
Open plan office																								
Intensive VDT use	Important	Important	Important	Very important	Important	Important	Important	Very important	Important		Very important	Important	Very important		Important		Somewhat important	(14,15)	Important	D	Very important	B		
Open plan office																								
Intermittent VDT use	Important	Important	Somewhat important	Very important	Important	Important	Important	Very important	Important		Very important	Important	Important		Important		Somewhat important	(14,15)	Important	E	Important	B		
Private office	Important	Important	Very important	Very important	Important	Important	Important	Very important	Important		Very important	Important	Important		Important				Important	E	Important	B		

FIGURE 26-31 The IESNA Handbook Design Guide categorizes important aspects of the visual environment and lists illuminance data for a wide variety of visual tasks. (*Reprinted from the IESNA Lighting Handbook, 9th eds., with permission from the IESNA.*)

informative computer renderings of the luminance patterns across a space. The lighting designer should also evaluate the interior. Battery-powered emergency incandescent-lamp units are widely employed to light aisles and corridors of industrial buildings, and equipping a predetermined portion of fluorescent luminaires with emergency ballasts can provide the prescribed illuminance levels for persons to leave their workstations and follow planned routes of egress. Such ballasts utilize high-temperature nickel-cadmium batteries, which results in fluorescent-lamp light output of sufficient duration to meet life safety code and UL standards (Table 26-17). Models are listed for both magnetic and electronic ballasts, for linear and compact fluorescent lamps, as combination units for both normal and emergency operation. In larger buildings, the emergency power source may be a generator that powers a number of different circuits within the building.

See Sec. 26.5 (paragraph on HID lamp operation) for standby illumination when HID lamps are extinguished by severe voltage dips or momentary power failures.

Veiling Reflections. Substantial losses in task contrast and hence in visibility and visual performance can result when light sources are reflected in such subtly specular (shiny) visual tasks as typing on bond paper. The apparent *veil* that is cast over a task when a light source is reflected in it may be so subtle as to be undetectable by the eye. Many factors contribute to veiling reflections and each of them, individually, has long been known. The problem is to integrate the effects of these interrelated factors. The factors are the visual task and its specularity; the worker's orientation, location, and viewing angles; and the lighting-system layout and luminaire light distribution and polarization.

The following are guidelines for reducing veiling reflections:

1. Where possible, written or printed tasks should be on matte paper using nongloss inks. The use of glossy paper stock and hard pencils should be minimized.
2. Light-source positions on either side and behind the workers are preferred.
3. Where work positions can be determined, substantial gains can be made by not positioning lighting equipment in the general area above and forward of the occupant's position for horizontal tasks, or at a position where veiling reflections would result for tasks of other orientations.
4. The use of luminaires with specific distributions and polarization characteristics for reducing reflections should be considered.
5. Side lighting such as from windows is effective in reducing veiling reflections. From the standpoint of visual comfort, workers should be positioned so their line of sight is parallel to or away from windows—rather than facing them.
6. Any decision on a lighting installation should also include considerations of its efficiency and the visual comfort in the space.

VDT (Computer) Criteria. One of the most common visual tasks that presents special lighting considerations is the visual display terminal (VDT). Although the front of the VDT is self-illuminated and the alphanumeric characters are visible without external lighting, there are typically other paper-background seeing tasks whose visibility are dependant on general or localized lighting systems. In

TABLE 26-17 Illuminance Levels for Safety*

Hazards requiring visual detection	Slight		High	
Normal activity level†	Low	High	Low	High
Illumination level				
Footcandles	0.5	1.0	2.0	5.0
Lux	5.4	11.0	22.0	54.0

*Minimum illumination for safety of personnel, absolute minimum at any time and at any location on any plane where safety is related to seeing conditions.
†Special conditions may require different levels of illumination.

addition, overall environmental luminance considerations call for general (ambient) illumination, which can be provided from a variety of luminaire distributions, all the way from totally direct to totally indirect.

A major specialized consideration involves the location and luminance distribution of luminaires in spaces containing VDTs, and the geometric relationship of the VDT screen to these luminaires and to windows. Luminance ratios should be no more than 3:1 between paper-based visual tasks and an adjacent VDT screen or dark surface, and the same 3 to 1 ratio applies between the VDT screen and an adjacent dark surface. For adjacent light surfaces, the corresponding luminance ratio is 1:3. For remote dark surfaces, the maximum ratio is 10:1, and for remote lighter surfaces, 1:10. To control direct glare, and to largely avoid bothersome reflections in the VDT screen, the average luminance of direct luminaires in the lengthwise, crosswise, and 45° vertical photometric planes at 55° from nadir should preferably not exceed 850 cd/m^2, 350 cd/m^2 at 65°, and 175 cd/m^2 at 75° or more from nadir. The average luminance should never exceed 850, 350, and 175 cd/m^2 at 65°, 75°, and 85°, respectively. Movable task lighting should not exceed 50 fc on the horizontal work plane (see IESNA publication RP-1-04, American National Standard Practice for Office Lighting).

Indirect systems should not exceed 850 cd/m^2 of ceiling luminance at any angle, or more than 8:1 (preferably 4:1) in uniformity. Luminaires with both upward and downward light components should have luminances of 850 cd/m^2 or less at angles between 55° and 90° in the crosswise, lengthwise, and 45° vertical photometric planes.

Visual Comfort. Visual comfort may occur when there are no overly high luminances within a worker's visual field. High luminances can also distract and even reduce visibility. Luminaires and fenestrations which have luminances that are too high for the environment in which they are located will produce discomfort. Discomfort from direct glare is reduced by

1. Decreasing the luminance of lighting equipment or other sources of objectionable glare, such as windows and overhead skylights.
2. Diminishing the area of uncomfortable luminances (with glare zone luminance constant).
3. Increasing the angle between the source and line of sight.
4. Increasing the general luminance in the room.

A rating system based on the degree of freedom from discomfort glare in a proposed lighting installation is called visual comfort probability (VCP). This system is limited to use with direct-lighting fluorescent troffers and cannot be applied to HID luminaires, CFL downlights, and many other direct distribution luminaires. Evaluation is based on the following factors which influence subjective judgments of visual comfort: room size and shape; room-surface reflectances; illumination levels; luminaire type, size, luminance, maximum luminance, and light distribution; number of luminaires; luminance of the field of view; observer location and line of sight; and differences in individual glare sensitivity. Since each of these factors can vary considerably, a standard set of conditions has been established and used as a basis for VCP tables.

Direct discomfort glare should not be considered a problem in lighting installations if all three of the following conditions are satisfied:

1. The VCP is 70 or more.
2. The ratio of maximum to average luminaire luminance does not exceed 5:1 (preferably 3:1) at 45°, 55°, 65°, 75°, and 85° from nadir, crosswise and lengthwise.
3. Maximum luminaire luminances, crosswise and lengthwise, do not exceed the following values:

Angle above nadir, deg	Maximum luminance, (cd/m^2)
45	7710
55	5500
65	3860
75	2570
85	1695

Luminance Distribution. The luminance relationship of the various surfaces in the visual field is important. When the eyes scan a task, an adaptation level is established consisting primarily of the task luminance. As the eyes leave the task and look at an area of a different luminance, there is a sudden loss of sensitivity to see the contrast of the detail in the new area until the eyes can readapt after a short length of time. In order to see the detail of a visual task accurately and quickly, luminance ratios of appreciable areas in the visual environment should be kept low. Table 26-18 is an example of recommended luminance ratios for offices (ANSI/IES RP-1-04, American National *Standard Practice for Office Lighting*). To help achieve these ratios, room surface should be of high reflectance, as recommended for specific applications.

Energy Conservation. In any lighting design it is important to consider the following recommendations for the reduction of energy used for lighting. They are appropriate for designing new construction, in renovation work, and in operating and maintaining existing installations.

1. Design Lighting for Expected Activity (Appropriate Illuminance for Seeing Tasks, with Less Light in Surrounding Nonwork Areas). In some interiors (general offices, classrooms, certain production spaces, etc.) where work stations are in close proximity to each other, the task areas represent a high percentage of the total work plane (length times width), and therefore reasonably uniform task-level illuminance is appropriate throughout the space. In other cases, private offices, for example, the visual tasks may be critical and prolonged for only 20 to 30% of the work plane (such as the tops of the desk and credenza of the single occupant). The lighting designer, thus, has the opportunity to conserve energy by reducing illuminance in the nonwork portions of the room. Such energy savings are in addition to those received by lowered illuminance levels in hallways, other circulation areas, storage spaces, etc.

Open-plan offices often have movable furniture containing linear fluorescent task lighting, and low partitions which subdivide the space to give worker privacy. Luminaires mounted under overhead cabinets or shelves should be positioned if possible to avoid veiling reflections, and a convenient “on-off” switch should be provided. Such supplementary luminaires should be selected to provide illuminances throughout the task area which adequately augment levels of the general lighting system, since the contribution from the general lighting system may be reduced by the height of the shelves, cabinets, partitions, etc.

2. Design with More Effective Luminaires and Fenestration. The goal of both daylight and artificial illumination systems is to produce visual environments which result in rapid and accurate seeing. The goal is achieved in an energy-conserving manner by selecting luminance distributions and luminaire positions which enhance task visibility. For example, desks, drafting tables, and many other types of workstations involve paper-background tasks which use pencil, ink, typed, photocopied, or printed characters which have a specular component of reflection. Such tasks are best lighted from the side. With some acoustical ceiling-suspension systems, recessed fluorescent luminaires can be easily moved as work stations change, especially if the luminaires are powered by plug-in “flexible” wiring cables and connectors (Fig. 26-32). Luminaires for three or four fluorescent lamps can be wired to two wall switches for manual selection of 50 or 100% illuminance if two two-lamp ballasts are involved, and for 33, 67 or 100% if one single-lamp and one two-lamp ballast operate a three-lamp luminaire. Such switching arrangements conveniently permit reduced illuminance

TABLE 26-18 Recommended Luminance Ratios for Offices*

To achieve a comfortable balance in the office and limit the effects of transient adaptations and disability glare, it is desirable and practical to limit luminance ratios between areas of appreciable size from normal viewpoints as follows:

1–√3	between task and adjacent light surroundings
1–√3	between task and adjacent dark surroundings
1–√5	between task and more remote darker surfaces
1–√10	between task and more remote lighter surfaces

*These ratios, 3:1 or 10:1 are recommended as maximums; reductions are generally beneficial.

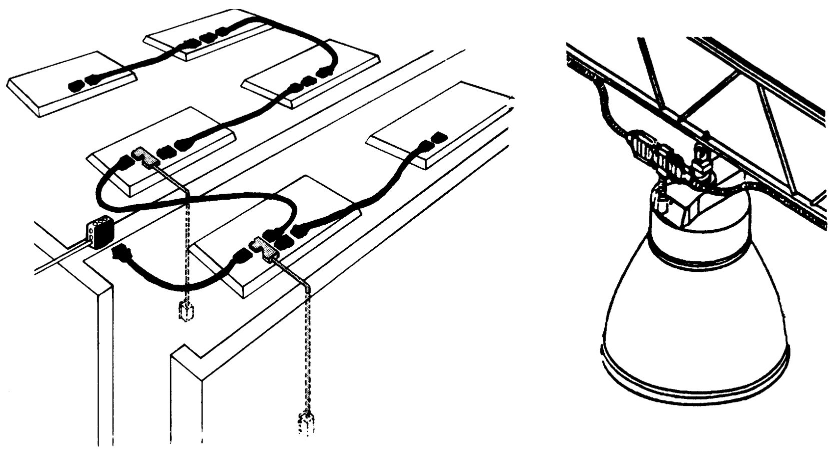


FIGURE 26-32 Flexible (plug-in) wiring systems for (a) commercial/institutional recessed and (b) industrial HID luminaires.

when it is adequate to supplement daylight in spaces with windows, and for the less demanding visual tasks during cleaning operations.

Daylight in areas which include exterior windows can be applied to save electric energy consumed by the electric lighting system. Because the quantity and quality of daylight varies greatly, windows require adjustable shades and blinds to control direct glare. Desks and other work stations should be oriented to minimize body shadows and to minimize veiling reflections (see IESNA publications RP-5, *Daylighting*).

Analysis of owning and operating charges can anticipate the present worth of future expenditures in a method known as "life cycle costing." Periods of 10 to 20 years are usually chosen as the "life" of a typical illumination system. The known or estimated time that future costs are expected to occur are converted to net present values. Such evaluations best show the economic benefits of more efficient, energy-conserving lighting systems.

3. Use Efficient Light Sources (Higher Lumen per Watt Output). Increased lamp efficacy obviously saves energy. Incandescent lamps have lower lumens per watt than any other family of general-service lamps. Also important are the annual hours that lighting is needed in a given space.* Table 26-12 lists many other characteristics that must be considered in selecting lamps. Some lumen per watt (lpw) guidelines:

Incandescent:

For equal life, lpw increases with wattage.

For comparable incandescent lamp options, increased life generally means lower lpw.

Krypton gas fill lamps give energy savings over lamp life.†

Tungsten-halogen capsules can give more lpw, or longer life, or both.

Undervoltage operation lowers lpw and increases life.‡

*Because of its small size and instant starting, and especially because of the low burning hours per year, incandescent lamps are the standard for broom closets even though their efficacy is low.

†Krypton-filled 34-, 52-, 67-, 90-, and 135-W longer-life lamps have lower efficacy than do 40-, 60-, 75-, 100-, and 150-W lamps.

‡For example, a 130-V 1000-h, 150-W argon-filled incandescent lamp operated at 120 V would have average life of approximately 1950 h, and produce initially approximately 2140 lm instead of its 2800-lm-rating if operated at 130 V.

Fluorescent:

32-W T-8 lamps are an obvious choice for 4-ft luminaires—greatly superior to 34-W cool white T-12 lamps in lumens per watt and CRI.

For given tube diameter and lamp current, long lamps have higher lpw than short lamps.

Longer burning hours per start increase lamp life.[§]

High CRI lamps using rare-earth phosphors have improved lpw as well as color rendition.

Compact fluorescent lamps can replace incandescent lamps in downlights, supplementary and adjustable task lights, and residential floor and table lamps.

System efficacy ratings should combine lamp, ballast, and luminaire performance.

High-intensity discharge:

If color is no factor, high-pressure sodium has highest lpw, long life, and good lumen maintenance. Metal halide provides good lpw and color. Since light output can decrease if operated in positions other than designated by the lamp manufacturer, a "tilt factor" is appropriate in design computations.

Standard Metal Halide lamps, in general, have noticeable problems with color consistency. Lamps shift to different shades of white and may not emit a uniform color of white light across all lamps in a system. Ceramic metal halide lamps and specially shaped arc-tube lamps help limit the color shift that occurs.

White high-pressure sodium lamps with 70 plus CRI ratings at 2800 K can be used with incandescent lamps (check for available wattages). 3200-K metal-halide lamps can be considered for retail and other commercial applications where warmer light-source color is desired.

4. Use More Efficient Luminaires. If a lighting system can deliver lumens to the work plane more efficiently, the wattage needed to achieve desired illuminance levels should be reduced, provided the luminaire intensity distribution is appropriate for the application being considered. Luminaire efficacy rating (LER) expresses, to the nearest whole number, the rated lumens per watt of a specific luminaire, equipped with specific lamps and ballasts.

The determination of LER considers luminaire efficiency to the nearest two-place decimal (three-place decimals may be used if desired), ballast factors rounded to the nearest two-place decimal, and input watts rounded to the nearest whole number. NEMA LE 5-2001 provides details on the calculation of LER[6], calling for photometry in accordance with IESNA LM-41, and defines the procedure for calculating ballast factor. Preceding the LER number in manufacturers' literature are two letters, indicating the luminaire category: FL = fluorescent recessed lensed, FP = fluorescent recessed parabolic, FW = fluorescent wraparounds, FS = fluorescent strip, and FI =

A similar NEMA Standard LE 5B-1998 for open (HO) and closed-bottom (HC) HID industrial luminaires[7], considered to have ballast factors of 1.0, gives minimum LERs for these luminaire types. Exempted are luminaires for HID lamps less than 150 W, special-application luminaires such as for hazardous areas, and luminaires using automatically switched quartz-lamp standby circuits.

Since rated ballast(s) input wattage is generated independently of photometric testing, existing photometric tests in compliance with IESNA LM-46 can be used to compute LER values. Because a variety of ballasts can be used with many luminaire designs, the specific ballast type must be reported in conjunction with the LER. See LE5B for further information.

1 Use Thermally Controlled Luminaires. Figure 26-12 shows that the ambient temperature near the luminaire, and the extent to which the heat produced by lamps and ballasts is retained, will have considerable effect on the light output. Tests show that maximum lamp efficacy occurs when the

[§]However, cathode improvements have greatly reduced the life penalty of shorter operating cycles. Combinations of lamp and energy cost may recommend turning off fluorescent lighting even for intervals as short as 10 min or less when the illumination is not needed.

lamp cold-spot temperature is approximately 104F.[8] A four-lamp 40-W troffer 2 ft wide and 4 ft long with a prismatic enclosure can have temperatures up to 120 to 130F inside the lamp compartment, which can account for approximately a 15% reduction in light output. By passing room air through the lamp compartment (called "heat transfer" or "heat exhaust"), a thermal condition more favorable to light production can be achieved, and less of the lighting heat enters the occupied space.

2 Use Lighter Finishes on Ceilings, Walls, Floors, and Furniture. IES recommended practices for various types of interiors list desirable reflectances for room surfaces and furniture. For commercial and institutional interiors these would be 80 to 90% for ceilings, 50 to 70% for walls, 40 to 70% for partitions, 25 to 45% for furniture, 25 to 45% for office machines and equipment, and 20 to 40% for floors. Similar values are suggested for industrial spaces (except not less than 20% for floors), realizing that they are much more difficult to control, and pointing out that some color can be stimulating. Because lower CRI lamps are widely used in factory areas, the colors should be selected while illuminated by the type of lamps under which they will be seen.

Turn Off Lights When Not Needed. Because time is a factor in the formula for electrical energy, programs to conserve energy should minimize the operation of lighting systems when spaces are unoccupied, even for a modest period appropriate for the *traffic* pattern. This is an owner benefit and energy-saving procedure, in addition to varying illuminances to supplement daylight, and to fit tasks that have higher or lower visual demands.

Building codes/standards such as ASHRAE/IESNA 90.1 mandate switches in each space and auto-shutoff. Clock switches can turn luminaires on and off, and can activate either or both of two circuits wired to three- and four-lamp luminaires, as previously mentioned. A central processing unit could provide the capability for preset, digital, and automated control, plus override arrangements when lighting is needed during normal off periods.

Occupancy sensors can also control lighting systems, with adjustable time intervals after infrared or ultrasonic beams sense the presence or absence of people in the space covered. Some take the place of wall light switches for use in private offices and other small spaces. Specifiers should compare features available from different manufacturers.

In warehouses, shipping docks, and many outdoor lighted areas there are numerous times when illuminances can be reduced to save energy if "full output" is provided almost instantly and automatically when needed. High-low components using occupancy detectors of the passive infrared or ultrasonic type can be positioned to sense motion in areas leading to warehouse aisles, for example, which could be operating as low as 30% of normal wattage. A coded signal would be transmitted without control wires to receivers located in metal halide or high-pressure sodium luminaires, resulting in near instant normal light level. A short interval after the motion ends, the luminaires would switch back automatically to the low illuminance level. The infrared occupancy detector/transmitters are smaller and work best outdoors and in low-temperature locations. Ultrasonic models have better sensitivity, can respond to the slightest motion, and are particularly effective over narrow warehouse aisles. General-purpose transmitters permit switches, relays, photocells, or time clocks to operate the high-low system.

8. Keep Lighting Equipment Clean and in Good Working Condition. Early in the design stage, lighting equipment should be selected that will be easy to maintain: to change lamps, to periodically clean reflecting and light-transmitting components, and to easily reach ballasts and other electrical auxiliaries. It is also important to mount luminaires so that maintenance personnel can have access for servicing, or provide means for lowering equipment to more convenient levels. The continuing effectiveness of lighting systems depends on good maintenance. In existing facilities, systematic reevaluations may reveal maintenance procedures which reduce the consumption of lighting energy.

Programs should include considerations of (1) group replacement of lamps at 70% to 80% of expected average lamp life and (2) scheduled luminaire cleanings that help the owner receive the light which the system was designed to provide. Both justify higher light loss factors to reduce the number of luminaires needed. Lighting maintenance companies in metropolitan areas can be engaged to perform relamping, cleaning, and repair functions, and to provide proper cleaning compounds, ladders, scaffolds, etc. In-house maintenance departments should develop and post instructions which cover procedures and frequency of different services, coordinated to minimize disruption of essential activities.

Lighting and the Thermal Environment. Effective building design requires a provision for efficient utilization of dissipation of lighting heat. The benefits of integrating building heat in lighting design are: (1) improved performance of the air-conditioning system, (2) more efficient handling of lighting heat, and (3) more efficient lamp performance if fluorescent lamps are the light source.

The control and removal of lighting heat before it enters the occupied space can reduce heat in that space, reduce air changes and fan horsepower, lower temperature differentials required in the space, enable a more economical cooling-coil selection because of the higher temperature differential across the coil, and reduce luminaire and ceiling temperature, thereby minimizing radiant effects.

The degree to which any of these benefits may be obtained depends on many variables such as the quantity of energy involved, the type of heat-transfer mechanism, the temperature difference between source and sink, and the velocity and quantity of air available for heat transfer.

REFERENCES ON LIGHTING DESIGN

1. RP-1-04, *American National Standard Practice for Office Lighting*, Illuminating Engineering Society of North America, www.iesna.org
2. RP-3-00, *Lighting for Educational Facilities*, Illuminating Engineering Society of North America, www.iesna.org
3. RP-7-01, *Lighting Industrial Facilities*, Illuminating Engineering Society of North America, www.iesna.org
4. RP-8-00, *Roadway Lighting*, Illuminating Engineering Society of North America, www.iesna.org
5. *IES Lighting Handbook*, 2000.
6. LE 5-2001, *Procedure for Determining Luminaire Efficacy Ratings for Fluorescent Luminaires*, National Electrical Manufacturers Association (NEMA), www.nema.org
7. LE 5B-1998, *Procedure for Determining Luminaire Efficacy Ratings for High-Intensity Discharge Industrial Luminaires*, National Electrical Manufacturers Association (NEMA), www.nema.org
8. Treado, S.J., *Lighting and HVAC, Lighting Design and Application*, July 1991, p. 18.

26.10 QUANTITY AND QUALITY OF ILLUMINATION

Lighting is provided so that people can perform visual tasks. Visual performance is a function of a number of fundamentally important factors. Some of these are the size of the object or detail to be seen, the contrast of the detail with its immediate background, the luminance of the object, the time available to see it, the luminance relation between the object and its surroundings, the visual capability of the human seeing machine, and the level of personal motivation.

The difficulty of the visual task is determined by the size and contrast of the task details. Other factors being equal, small tasks require more illuminance for a given level of visual performance than larger tasks. Similarly, low contrast between the task details and their background justifies higher illuminance levels. The IESNA procedure for selecting illuminance[5] employs seven generic types of interior activity, identified by the first seven letters of the alphabet, as described in Table 26-19, Letter categories A, B, and C, called for modest levels of general lighting throughout the spaces. Categories D, E, and F required minimum illuminances on the task, and G involves higher illuminances on more difficult tasks, obtained by a combination of general and local (supplementary) lighting.

In the past, modifications to these recommended illuminance levels were made based on characteristics of the space, task and age of the occupants. In the current method, the IESNA provides less guidance on when to increase or decrease the recommended values, but suggests that 1/3 increase or decrease to these values be applied if actual conditions or the occupant's age justifies such a change.

TABLE 26-19 Typical Recommended Illuminances

Area or activity	Illuminance category	Recommended illuminance, fc*
Industrial assembly or inspection		
Simple	D	30
Moderately difficult	E	50
Difficult	F	100
Exacting	G	300‡
Machine shops		
Rough bench or machine work	D	30
Medium bench or machine work, ordinary automatic machines, rough grinding, medium buffing and polishing	E	50
Fine bench or machine work, fine automatic machines, medium grinding, fine buffing and polishing	G	300
Extrafine bench or machine work, grinding, fine work	G	300‡
Materials handling		
Wrapping, packing, labeling	D	30
Picking stock, classifying	D	30‡
Loading, inside truck bodies and freight cars	C	10
Reading copied tasks		
Thermal copy, poor copy	F	75§
Xerograph	D	30
Xerography, third generation and greater	E	50
Reading handwritten tasks		
#3 pencil and softer leads	E	50§
#4 pencil and harder leads	F	75§
Ball-point pen	D	30§
Felt-tip pen	D	30
Handwritten carbon copies	E	50
Whitelooards	D	30
Chalkboards	E	50§
Reading printed tasks		
6-point type	E	50§
8- and 10-point type	D	30§
Glossy magazines	D	30§
Maps	E	50
Newsprint	D	30
Typed originals	D	30
Typed second carbon and later	E	50
Telephone books	E	50

*See 9th or later edition of *IES Lighting Handbook* for complete listing. Multiply footcandle values listed above by 10 for corresponding rounded-off lux values (1 footcandle = 10.76 lux).

†Use information in Table 26-26 to determine whether "normal" or "difficult" task column is appropriate.

‡Obtained by a combination of general and supplementary lighting.

§Subject to veiling reflections. Glossy surfaces are especially subject to veiling reflections, and it may be necessary to shield the task, reorient it, or select a lighting system that avoids veiling reflections.

26.11 CALCULATING MAINTAINED ILLUMINANCE

Maintained illuminance is the illuminance that will be provided by a lighting system at some future point in time, and generally considers a variety of light-loss factors. Lighting systems are generally designed so that the maintained illuminance prior to cleaning, relamping, etc. is near the target illuminance level. A new and clean lighting system, therefore, will typically provide work plane illuminance levels above the target level to account for the losses that occur over time.

Light-Loss Factor. In calculating maintained illuminance (the lowest level before maintenance procedures are instituted) a light-loss factor (LLF) is used to take into account losses in light output due to temperature and voltage variations, ballast factors, lamp position (tilt) factors, equipment operating factors, thermal factors, dirt accumulation on luminaire and room surfaces, lamp depreciation, maintenance procedures, and atmospheric conditions. The effect of these light losses can be seen in the example shown in Fig. 26-33. It is important that all causes of light loss be investigated and maintenance procedures be established in the lighting-design stage to assure adequate illumination is provided in an economical manner without wasting energy.

LLF Determination. The LLF (formerly called maintenance factor) is mainly the product of the ballast factor (BF), lamp lumen depreciation factor (LLD), the luminaire dirt depreciation factor (LDD), the room-surface dirt depreciation factor (RSDD), and lamp burnout factor (LBO). Other factors such as ambient temperature factor (for fluorescent) and voltage-to-luminaire factor (for incandescent) may also be applied. Data on LLD are available from lamp manufacturers. LBO is the

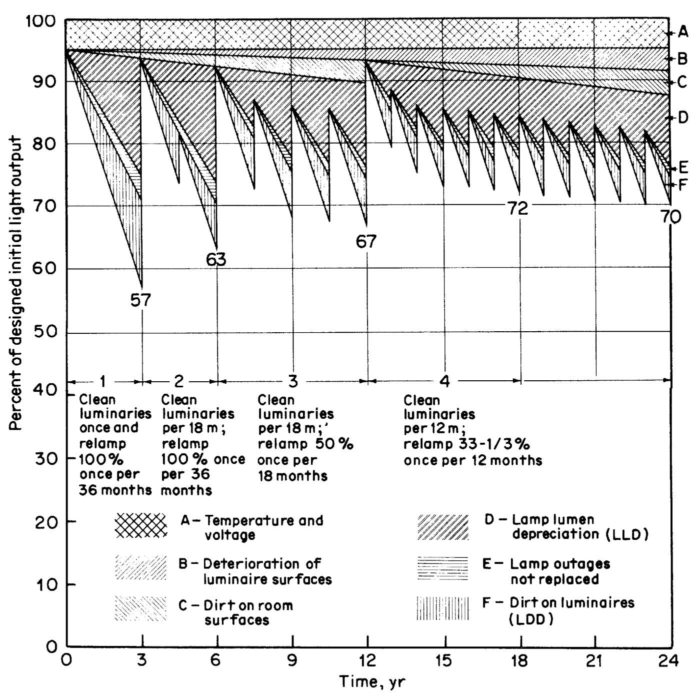


FIGURE 26-33 Effect of light loss on illumination level. Example shown is for enclosed surface-mounted luminaires with 40-W fluorescent lamps, operated 10 h a day, 5 days per week, 2600 h per year. Four maintenance system are shown for comparison purposes.

ratio of lamps remaining lighted to the total in the original installation, if burnouts are not replaced immediately. LDD is obtained from curves, such as shown in Fig. 26-34, a knowledge of the amount of dirt in the room atmosphere, a knowledge of the length of time between luminaire cleanings, and the luminaire maintenance category, a Roman numeral from I to VI, which is based on the luminaire type (see Table 26-20). RSDD is obtained from a curve and table as shown in Fig. 26-35, where room size, luminaire distribution, room atmosphere, and cleaning cycle are known.

Example of calculation procedure. Given a direct type of luminaire of category IV, with 32-W fluorescent lamps (LLD = 0.92. Room atmosphere is very clean and cleaning is performed every 18 months. Room-cavity ratio is 2. Lamps are replaced at burnout.

BF = 0.94 (from ballast or luminaire manufacturer)

LLD = 0.92 (mean lumens/initial lumens form manufacturer's lamp tables)

LDD from Fig. 26-34 5 0.91

RSDD from Fig. 26-35 = 0.98

LBO = 1

$\mathrm{LLF} = 0.94 \times 0.92 \times 0.91 \times 0.98 \times 1.0 = 0.77$

26.12 CALCULATION OF AVERAGE ILLUMINANCE

General Lighting. The design of general lighting systems is governed by room dimensions, structural features, reflectances of room surfaces, mounting height of the luminaires, and the distribution and maintenance characteristics of the luminaire. The choice of the luminaire depends on the service to which it is to be put, which assumes a certain experience in selection, or other aids such as manufacturers' data, which assist the designer in making a selection appropriate from the standpoint of freedom from glare, efficiency, decorative value, and economy. The ultimate "luminance pattern" of the room is an important factor in the design.

The beginning concept of general lighting design is that of delivering a specified average illuminance level to a horizontal plane in a room. The light generated by the lamps in such a system is variously affected and considerably reduced by reflection, diffusion, and absorption as it impinges on reflectors and transmitting media in the luminaires and on ceilings, wall, floors, and on the objects in the room.

The Lumen (Zonal-Cavity) Method.[1] The lumen method is used in calculating the illuminance that represents the average of all points on the work plane in an interior. It is based on the definition of illuminance as luminous flux per unit area, or

where luminous flux is expressed in lumens. If the area is in square feet, the illuminance is in footcandles (lumens per square foot); if the area is in square meters, the illuminance is in lux (lumens per square meter).

Because not all the lamp lumens will reach the work plane owing to losses in the luminaire and at the room surfaces, they must be multiplied by a coefficient of utilization which represents the portion that reaches the work plane.

Since, the design objective is usually to maintain a minimum value of illuminance, factors must be applied to account for the estimated depreciation in lamp lumens, the estimated losses from dirt collection on the luminaire surfaces (including lamps), etc., that occur over time, or due to system components.

The zonal-cavity method considers the actual room as being made up of a ceiling cavity above the luminaires, a floor cavity beneath the work plane, and a room cavity located between the two (see Fig. 26-36).

In the general case, all these cavities are present. In the case of recessed or surface-mounted luminaires, the ceiling cavity is simply the ceiling. When the illuminance on the floor is to be determined, the floor cavity becomes the floor.

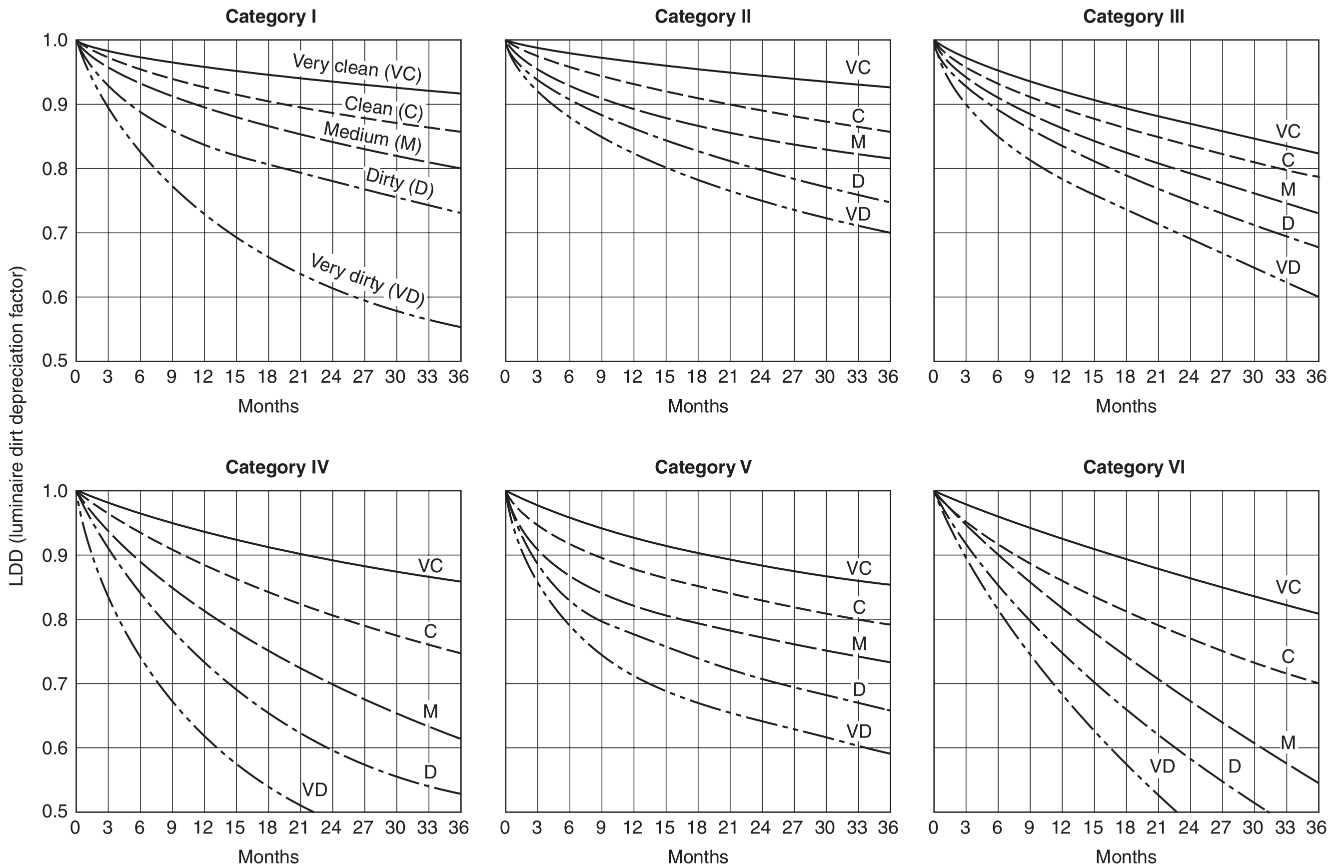


FIGURE 26-34 Luminaire dirt depreciation (LDD) factors for the six luminaire categories and for 5 levels of atmosphere dirtiness. (*Reprinted from the IESNA Lighting Handbook, 9th ed., with permission from the IESNA.*)

TABLE 26-20 A General interpretation of the IESNA Luminaire Maintenance Categories for Luminaires

Maintenance category	Types of luminaires included in this category
I	Bare lamps, bare lamp strips, and globes
II	Open pendant luminaires with >15% uplight through top openings
III	Open pendant luminaires with <15% uplight through top openings
IV	Luminaires with an open bottom and closed top, such as a parabolic troffer or open downlight
V	Luminaires that are completely enclosed, typically with lenses
VI	Uplights with an open top and solid bottom

It is now possible to calculate numerical relationships called *cavity ratios*, which may be used to determine effective reflectance of the floor and ceiling and then to find the coefficient of utilization. The basic steps in the calculation of any average illuminance are as follows:

1. Determine cavity ratios for three cavities shown in Fig. 26-36 as follows:

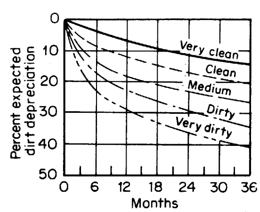
Room-cavity ratio, $\text{RCR} = 5 \times h_{RC}\,(L + W)/(L \times W)$

Ceiling-cavity ratio, $\text{CCR} = 5 \times h_{CC}\,(L + W)/(L \times W) = \text{RCR}\; h_{cc}/h_{Rc}$

Floor-cavity ratio, $\text{FCR} = 5 \times h_{FC}\,(L + W)/(L \times W) = \text{RCR}\; h_{F_c}/h_{Rc}$

Where h_{RC} = height of room between luminaire plane and work plane; h_{CC} = distance from luminaire plane to ceiling; h_{FC} = height of work plane above floor; L = room length; and W = room width.

2. Obtain the *effective floor-cavity reflectance* (ρfc) for the combination of floor and wall reflectances to be employed from Table 26-22.



Percent expected dirt depreciation	Luminaire distribution type: Direct 10	Direct 20	Direct 30	Direct 40	Semidirect 10	Semidirect 20	Semidirect 30	Semidirect 40	Direct-indirect 10	Direct-indirect 20	Direct-indirect 30	Direct-indirect 40	Semi-indirect 10	Semi-indirect 20	Semi-indirect 30	Semi-indirect 40	Indirect 10	Indirect 20	Indirect 30	Indirect 40
Room-cavity ratio																				
1	.98	.96	.94	.92	.97	.92	.89	.84	.94	.87	.80	.76	.94	.87	.80	.73	.90	.80	.70	.60
2	.98	.96	.94	.92	.96	.92	.88	.83	.94	.87	.80	.75	.94	.87	.79	.72	.90	.80	.69	.59
3	.98	.95	.93	.90	.96	.91	.87	.82	.94	.86	.79	.74	.94	.86	.78	.71	.90	.79	.68	.58
4	.97	.95	.92	.90	.95	.90	.85	.80	.94	.86	.79	.73	.94	.86	.78	.70	.89	.78	.67	.56
5	.97	.94	.91	.89	.94	.90	.84	.79	.93	.86	.78	.72	.93	.86	.77	.69	.89	.78	.66	.55
6	.97	.94	.91	.88	.94	.89	.83	.78	.93	.85	.78	.71	.93	.85	.76	.68	.89	.77	.66	.54
7	.97	.94	.90	.87	.93	.88	.82	.77	.93	.84	.77	.70	.93	.84	.76	.68	.89	.76	.65	.53
8	.96	.93	.89	.86	.93	.87	.81	.75	.93	.84	.76	.69	.93	.84	.76	.68	.88	.76	.64	.52
9	.96	.92	.88	.85	.93	.87	.80	.74	.93	.84	.76	.68	.93	.84	.75	.67	.88	.75	.63	.51
10	.90	.92	.87	.83	.93	.86	.79	.72	.93	.84	.75	.67	.92	.83	.75	.67	.88	.75	.62	.50

FIGURE 26-35 Room-surface dirt depreciation (RSDD) factors, based on space dirt level, room surface cleaning interval, and CIE luminaire distribution type.

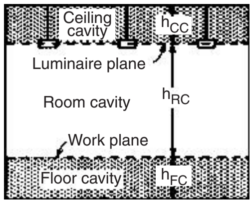


FIGURE 26-36 Room cavities.

3. Obtain the *effective ceiling-cavity reflectance* (ρcc) for the combination of ceiling and wall reflectance to be employed from Table 26-23. Note that, for surface-mounted or recessed luminaires, CCR = 0 and the ceiling reflectance may be used as the effective cavity reflectance.

4. Obtain the coefficient of utilization for the luminaire for a 20% effective floor-cavity-reflectance condition (tables like 26-21), interpolating between tabulated values as required to match room size and ceiling- and wall-reflectance combinations.

5. If the *effective floor-cavity reflectance* (ρfc) obtained in step 3 differs significantly from 20%, obtain a multiplier form Table 26-23 or Table 26-24. Multiply the coefficient of utilization by this multiplier.

6. Determine average maintained illuminance by the following formula:

$$E_{ave} = \frac{(\text{rated lamp lumens / luminaire}) \times \text{CU} \times \text{LLF*}}{\text{area / luminaire}}$$

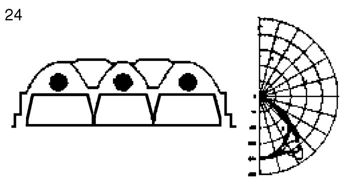
If initial illuminance is desired, use a LLF factor that considers only the LLF components that will occur when the system is new, such as the ballast factor. When the desired average maintained illumination is known, the formula can more conveniently be expressed as

$$\text{Number of luminaires required} = \frac{E_{desired} \times (\text{room area})}{(\text{rated lamp lumens / luminaire}) \times \text{CU} \times \text{LLF*}}$$

$$\text{Area / luminaire} = \frac{(\text{rated lamp lumens / luminaire}) \times \text{CU} \times \text{LLF*}}{E_{desired}}$$

This area per luminaire when divided by the luminaire length gives the approximate spacing between continuous rows, or it may be divided into the total room area to determine the number of

TABLE 26-21 Coefficients of Utilization of a Troffer Designed for Spaces with VDTs



2 × 4.3-Lamp parabolic troffer with 3" semi-spac. louvers, 18 cells

ρcc →	80			70			50			30			10			0
ρw →	70	50	30	70	50	30	50	30	10	50	30	10	50	30	10	0
RCR	EFF = 72.7%			% DN = 100			% UP = 0			Lamp = (3) F32T8 SC (along, across, 45") = 1.3, 1.6, 1.6						
0	0.87	0.87	0.87	0.85	0.85	0.85	0.81	0.81	0.81	0.77	0.77	0.77	0.74	0.74	0.74	0.73
1	0.81	0.78	0.76	0.79	0.77	0.74	0.74	0.72	0.70	0.71	0.69	0.68	0.68	0.67	0.66	0.65
2	0.75	0.70	0.66	0.73	0.69	0.65	0.66	0.63	0.61	0.64	0.61	0.59	0.62	0.60	0.58	0.57
3	0.69	0.63	0.58	0.68	0.62	0.57	0.60	0.56	0.52	0.58	0.54	0.52	0.56	0.53	0.51	0.49
4	0.64	0.56	0.51	0.62	0.55	0.50	0.54	0.49	0.46	0.52	0.48	0.45	0.51	0.47	0.44	0.43
5	0.59	0.51	0.45	0.58	0.50	0.44	0.48	0.44	0.40	0.47	0.43	0.40	0.46	0.42	0.39	0.38
6	0.55	0.46	0.40	0.53	0.45	0.40	0.44	0.39	0.35	0.43	0.38	0.35	0.42	0.38	0.35	0.33
7	0.51	0.42	0.36	0.50	0.41	0.36	0.40	0.35	0.31	0.39	0.35	0.31	0.38	0.34	0.31	0.30
8	0.47	0.38	0.32	0.46	0.38	0.32	0.37	0.32	0.28	0.36	0.31	0.28	0.35	0.31	0.28	0.27
9	0.44	0.35	0.29	0.43	0.35	0.29	0.34	0.29	0.25	0.33	0.29	0.25	0.32	0.28	0.25	0.24
10	0.41	0.32	0.27	0.40	0.32	0.27	0.31	0.26	0.23	0.31	0.26	0.23	0.30	0.26	0.23	0.22

* Include ballast and tilt factors in LLF if appropriate.

TABLE 26-22 Percent Effective Ceiling- or Floor-Cavity Reflectance for Various Reflectance Combinations*

% base reflectance†	90				80				70			50			30				10		
% wall reflectance	90	70	50	30	80	70	50	30	70	50	30	70	50	30	65	50	30	10	50	30	10
Cavity ratio																					
0	90	90	90	90	80	80	80	80	70	70	70	49	50	50	30	30	30	30	10	10	10
0.1	90	89	88	87	79	79	78	78	69	69	68	49	49	48	30	30	29	29	10	10	10
0.2	89	88	86	85	79	78	77	76	68	67	66	49	48	47	30	29	29	28	10	10	9
0.3	89	87	85	83	78	77	75	74	68	66	64	49	47	46	30	29	28	27	10	10	9
0.4	88	86	83	81	78	76	74	72	67	65	63	48	46	45	30	29	27	26	11	10	9
0.5	88	85	81	78	77	75	73	70	66	64	61	48	46	44	29	28	27	25	11	10	9
0.6	88	84	80	76	77	75	71	68	65	62	59	47	45	43	29	28	26	25	11	10	9
0.7	88	83	78	74	76	74	70	66	65	61	58	47	44	42	29	28	26	24	11	10	8
0.8	87	82	77	73	75	73	69	65	64	60	56	47	43	41	29	27	25	23	11	10	8
0.9	87	81	76	71	75	72	68	63	63	59	55	46	43	40	29	27	25	22	11	9	8
1.0	86	80	74	69	74	71	66	61	63	58	53	46	42	39	29	27	24	22	11	9	8
1.1	86	79	73	67	74	71	65	60	62	57	52	46	41	38	29	26	24	21	11	9	8
1.2	86	78	72	65	73	70	64	58	61	56	50	45	41	37	29	26	23	20	12	9	7
1.3	85	78	70	64	73	69	63	57	61	55	49	45	40	36	29	26	23	20	12	9	7
1.4	85	77	69	62	72	68	62	55	60	54	48	45	40	35	28	26	22	19	12	9	7
1.5	85	76	68	61	72	68	61	54	59	53	47	44	39	34	28	25	22	18	12	9	7
1.6	85	75	66	59	71	67	60	53	59	53	45	44	39	33	28	25	21	18	12	9	7
1.7	84	74	65	58	71	66	59	52	58	51	44	44	38	32	28	25	21	17	12	9	7
1.8	84	73	64	56	70	65	58	50	57	50	43	43	37	32	28	25	21	17	12	9	6
1.9	84	73	63	55	70	65	57	49	57	49	42	43	37	31	28	25	20	16	12	9	6
2.0	83	72	62	53	69	64	56	48	56	48	41	43	37	30	28	24	20	16	12	9	6
2.1	83	71	61	52	69	63	55	47	56	47	40	43	36	29	28	24	20	16	13	9	6
2.2	83	70	60	51	68	63	54	45	55	46	39	42	36	29	28	24	19	15	13	9	6
2.3	83	69	59	50	68	62	53	44	54	46	38	42	35	28	28	24	19	15	13	9	6
2.4	82	68	58	48	67	61	52	43	54	45	37	42	35	27	28	24	19	14	13	9	6
2.5	82	68	57	47	67	61	51	42	53	44	36	41	34	27	27	23	18	14	13	9	6

TABLE 26-22 Percent Effective Ceiling- or Floor-Cavity Reflectance for Various Reflectance Combinations* (*Continued*)

% base reflectance†	90				80				70			50			30				10		
% wall reflectance	90	70	50	30	80	70	50	30	70	50	30	70	50	30	65	50	30	10	50	30	10
2.6	82	67	56	46	66	60	50	41	53	43	35	41	34	26	27	23	18	13	13	9	5
2.7	82	66	55	45	66	60	49	40	52	43	34	41	33	26	27	23	18	13	13	9	5
2.8	81	66	54	44	66	59	48	39	52	42	33	41	33	25	27	23	18	13	13	9	5
2.9	81	65	53	43	65	58	48	38	51	41	33	40	33	25	27	23	17	12	13	9	5
3.0	81	64	52	42	65	58	47	38	51	40	32	40	32	24	27	22	17	12	13	8	5
3.1	80	64	51	41	64	57	46	37	50	40	31	40	32	24	27	22	17	12	13	8	5
3.2	80	63	50	40	64	57	45	36	50	39	30	40	31	23	27	22	16	11	13	8	5
3.3	80	62	49	39	64	56	44	35	49	39	30	39	31	23	27	22	16	11	13	8	5
3.4	80	62	48	38	63	56	44	34	49	38	29	39	31	22	27	22	16	11	13	8	5
3.5	79	61	48	37	63	55	43	33	48	38	29	39	30	22	26	22	16	11	13	8	5
3.6	79	60	47	36	62	54	42	33	48	37	28	39	30	21	26	21	15	10	13	8	5
3.7	79	60	46	35	62	54	42	32	48	37	27	38	30	21	26	21	15	10	13	8	4
3.8	79	59	45	35	62	53	41	31	47	36	27	38	29	21	26	21	15	10	13	8	4
3.9	78	59	45	34	61	53	40	30	47	36	26	38	29	20	26	21	15	10	13	8	4
4.0	78	58	44	33	61	52	40	30	46	35	26	38	29	20	26	21	15	9	13	8	4
4.1	78	57	43	32	60	52	39	29	46	35	25	37	28	20	26	21	14	9	13	8	4
4.2	78	57	43	32	60	51	39	29	46	34	25	37	28	19	26	20	14	9	13	8	4
4.3	78	56	42	31	60	51	38	28	45	34	25	37	28	19	26	20	14	9	13	8	4
4.4	77	56	41	30	59	51	38	28	45	34	24	37	27	19	26	20	14	8	13	8	4
4.5	77	55	41	30	59	50	37	27	45	33	24	37	27	19	25	20	14	8	14	8	4
4.6	77	55	40	29	59	50	37	26	44	33	24	36	27	18	25	20	14	8	14	8	4
4.7	77	54	40	29	58	49	36	26	44	33	23	36	26	18	25	20	13	8	14	8	4
4.8	76	54	39	28	58	49	36	25	44	32	23	36	26	18	25	19	13	8	14	8	4
4.9	76	53	38	28	58	49	35	25	44	32	23	36	26	18	25	19	13	7	14	8	4
5.0	76	53	38	27	57	48	35	25	43	32	22	36	26	17	25	19	13	7	14	8	4

*Tabular values based on 1.6 length-to-width ratio.
†Ceiling, floor or floor of cavity.

TABLE 26–23 Multiplying Factors for 10% Effective Floor-Cavity Reflectance (20% = 1.00)

Effective ceiling cavity reflectance, ρ_{cc}	80				70				50			30			10		
% wall reflectance, ρ_w	70	50	30	10	70	50	30	10	50	30	10	50	30	10	50	30	10
Room cavity ratio:																	
1	0.923	0.929	0.935	0.940	0.933	0.939	0.943	0.948	0.956	0.960	0.963	0.973	0.976	0.979	0.989	0.991	0.993
2	0.931	0.942	0.950	0.958	0.940	0.949	0.957	0.963	0.962	0.968	0.974	0.976	0.980	0.985	0.988	0.991	0.995
3	0.939	0.951	0.961	0.969	0.945	0.957	0.966	0.973	0.967	0.975	0.981	0.978	0.983	0.988	0.988	0.992	0.996
4	0.944	0.958	0.969	0.978	0.950	0.963	0.973	0.980	0.972	0.980	0.986	0.980	0.986	0.991	0.987	0.992	0.996
5	0.949	0.964	0.976	0.983	0.954	0.968	0.978	0.985	0.975	0.983	0.989	0.981	0.988	0.993	0.987	0.992	0.997
6	0.953	0.969	0.980	0.986	0.958	0.972	0.982	0.989	0.977	0.985	0.992	0.982	0.989	0.995	0.987	0.993	0.997
7	0.957	0.973	0.983	0.991	0.961	0.975	0.985	0.991	0.979	0.987	0.994	0.983	0.990	0.996	0.987	0.993	0.998
8	0.960	0.976	0.986	0.993	0.963	0.977	0.987	0.993	0.981	0.988	0.995	0.984	0.991	0.997	0.987	0.994	0.998
9	0.963	0.978	0.987	0.994	0.965	0.979	0.989	0.994	0.983	0.990	0.996	0.985	0.992	0.998	0.988	0.994	0.999
10	0.965	0.980	0.989	0.995	0.967	0.981	0.990	0.995	0.987	0.991	0.997	0.986	0.993	0.998	0.988	0.994	0.999

TABLE 26–24 Multiplying Factors for 30% Effective Floor-Cavity Reflectance (20% = 1.00)

% effective ceiling cavity reflectance, ρ_{cc}	80				70				50			30			10		
% wall reflectance, ρ_w	70	50	30	10	70	50	30	10	50	30	10	50	30	10	50	30	10
Room cavity ratio:																	
1	1.092	1.082	1.075	1.068	1.077	1.070	1.064	1.059	1.049	1.044	1.040	1.028	1.026	1.023	1.012	1.010	1.008
2	1.079	1.066	1.055	1.047	1.068	1.057	1.048	1.039	1.041	1.033	1.027	1.026	1.021	1.017	1.013	1.010	1.006
3	1.070	1.054	1.042	1.033	1.061	1.048	1.037	1.028	1.034	1.027	1.020	1.024	1.017	1.012	1.014	1.009	1.005
4	1.062	1.045	1.033	1.024	1.055	1.040	1.029	1.021	1.030	1.022	1.015	1.022	1.015	1.010	1.014	1.009	1.004
5	1.056	1.038	1.026	1.018	1.050	1.034	1.024	1.015	1.027	1.018	1.012	1.020	1.013	1.008	1.014	1.009	1.004
6	1.062	1.033	1.021	1.014	1.047	1.030	1.020	1.012	1.024	1.015	1.009	1.019	1.012	1.006	1.014	1.008	1.003
7	1.047	1.029	1.018	1.011	1.043	1.026	1.017	1.009	1.022	1.013	1.007	1.018	1.010	1.005	1.014	1.008	1.003
8	1.044	1.026	1.015	1.009	1.040	1.024	1.015	1.007	1.020	1.012	1.006	1.017	1.009	1.004	1.013	1.007	1.003
9	1.040	1.024	1.014	1.007	1.037	1.022	1.014	1.006	1.019	1.011	1.005	1.016	1.009	1.004	1.013	1.007	1.002
10	1.037	1.022	1.012	1.006	1.034	1.020	1.012	1.005	1.017	1.010	1.004	1.015	1.009	1.003	1.013	1.007	1.002

luminaires required. For direct and semidirect luminaires, spacing between columns and rows should not exceed the maximum spacing as determined by the spacing criterion values for the luminaire.

Example. A room is 28 ft wide and 32 ft long and has a 10-ft ceiling height.

Reflectances are: ceiling 80%, walls 50%, floor 10%. A recessed four-lamp fluorescent luminaire is to be used. Work plane is 2 ft 0 in. Find the coefficient of utilization.

1. In Table 26-21, lookup *effective cavity reflectances* for ceiling and floor cavities. ρcc for the ceiling cavity will be 80%, while ρfc for the floor cavity will be 11%.
2. With the room-cavity ratio RCR known, it is now possible to find the coefficient of utilization for the luminaire in a room having an RCR of 2.7 and effective reflectances as follows (assume luminaire is type described in Table 26-21):

$$\rho cc = 80\% \ \rho w = 50\% \ \rho fc = 20\%$$

Thus CU = 0.65. Note that this is for an effective floor reflectance of 20%, while the actual effective reflectance of the floor ρfc is 11%. To correct for this, locate the appropriate multiplier in Table 26-24 for the RCR already calculated (2.7). It is 0.948 and is found by interpolating between the numbers for 80 ρcc and between RCRs of 2.0 and 3.0. Then

$$\text{CU final} = 0.65 \times 0.948 \times 0.62$$

The average illuminance across the space can now be calculated if we know the lamp lumen rating, and the light-loss factor.

$$E_{ave} = \frac{(\text{number of luminaires}) \times (\text{rated lamp lumens / luminaire}) \times \text{CU} \times \text{LLF}}{\text{room area}}$$

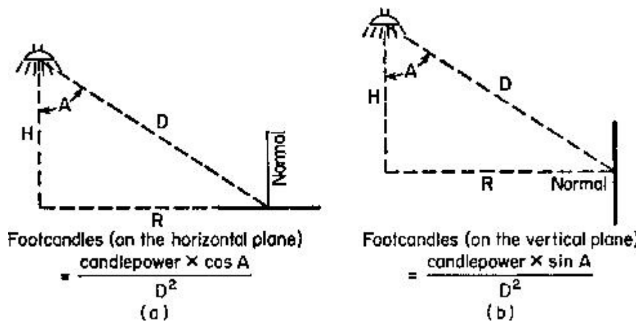
Simplified Design Methods. Buildings may have many individual spaces of different dimensions. If all of the above steps were used to determine the number of luminaires (and/or the number of lamps per luminaire) for each space, the procedure might be too time-consuming or unnecessary accurate or both. Engineers may wish, therefore, to use simplified performance data provided by the luminaire manufacturer in technical portions of catalogs, etc.

One popular method is based on specific illuminance levels for a small group of room proportions (ratios of room width to room or mounting height), and for assumed room-surface reflectances. By using typical light-loss factors, the approximate square footage per luminaire can be calculated using the formula under item 6 of the basic steps in calculating average illuminance.

26.13 CALCULATION OF ILLUMINANCE AT A POINT

Point Calculation Methods. The calculation of the illuminance at a point, whether on a horizontal, a vertical, or an inclined plane consists of two parts—the direct component and the reflected component. The total of these two components is the illuminance at the point in question.

Direct Component Illuminance. Most methods are based on the application of the *inverse-square law* ($E = I/D^2$, where I equals the luminous intensity expressed in candelas obtained from the candlepower distribution curve for the luminaire). In applying this formula to a horizontal surface, the illuminance at any point P is equal to the candlepower directed toward P multiplied by the cosine of the angle A and divided by the square of the distance D from the luminaire [$E = (I \times \cos A)/D^2$]. In this case A is the angle between the axis of the luminaire and a line from the light center to point P.



Reflected Component Illuminance. The reflected component consists of light that strikes the work plane, or any other point of interest after reflection from one or more of the room surfaces. In direct lighting systems and large rooms, this component may be small (less than 20% of the total illuminance at a point), while for indirect lighting systems, all of the light striking the work plane may be reflected. Illuminance calculated with the lumen method includes both the average direct and reflected illuminance on the work plane. For more detailed point-by-point calculations, advanced computer software often apply algorithms that subdivide the room surfaces into smaller patches for analysis of the reflected contribution, providing grids of points at which the illuminance is computed due to both direct and reflected light. Many of these computer programs also provide renderings of how an architectural space might appear. These tools are useful for analyzing lighting systems that do not provide uniform illuminance over the entire space, or to address conditions over a specific area or task location. A survey of lighting software is regularly published in the IESNA's periodical *Lighting Design + Application*.[1]

REFERENCE ON LIGHTING CALCULATIONS

1. Rea, M.S., Ed., *IES Lighting Handbook*, Chapter 9. IESNA 2000.

26.14 FLOODLIGHTING DESIGN AND PROCEDURE

Beam-Lumen Method. All floodlighting-design methods include certain approximations, based on experience. In floodlighting systems containing a large number of luminaires, a detailed computer study of point source calculations with luminaire locations and aiming positions is usually required. A procedure which is useful for designing simpler systems is called the "beam-lumen" method. This method requires the solution of the two formulas (A and B), as discussed in the following paragraphs, and the coordination of the results.

In many locations in which floodlighting is proposed, there are some basic dimensions that can be assumed to be already fixed. For example, in ground-area floodlighting, the designer is usually able to locate points where the equipment should logically be placed, such as on poles/towers or other physical structures. These locations establish the approximate perpendicular distance D from the floodlight to the plane of the surface to be lighted and the average aiming angles. They also guide the choice of floodlight type—narrow, medium, or broad-beam. In like manner the choice of equipment for lighting vertical surfaces can be obtained by taking D as the horizontal distance from the luminaire to the plane in which the vertical surface is located.

The average aiming angle is measured from the perpendicular to the beam-axis line (Fig. 26-37). In a perimeter system in which the floodlights are mounted along or beyond the perimeter of an area, they will, of course, be aimed at various angles, but the average aiming angle used in computation is measured between the perpendicular and the centerline of the area to be lighted. When floodlights are on poles along the centerline of an area, the average aiming angle is measured between the pole (perpendicular) and a point halfway to the boundary (one-fourth of the width of the total area).

Design Formulas and Procedure

$$\text{Formula A} = \frac{(\text{area lighted}) \times (\text{coverage factor})}{\text{floodlights needed for coverage} \times \text{beam-spot area}}$$

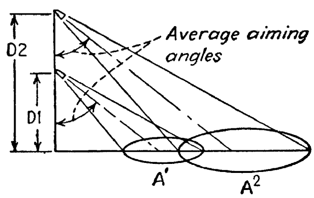


FIGURE 26-37 Spot areas (for the same beam spread and aiming angle) vary as the square of the distance *D*. Spot length *L* and spot width *W* vary as the distance *D*. The spot area may be determined from $\pi LW/4$.

In this formula, the area to be lighted may be either a horizontal surface or a vertical one.

The coverage factor indicates the minimum number of directions from which each point in the area should be lighted, depending on the use of the area. A coverage factor of 1 is acceptable in some applications, although in such systems one or two lamp burnouts might temporarily leave large dark patches. Coverage factors greater than 1 therefore add desirable safety factors. For example, a coverage factor of 2 is necessary for parking spaces and for protective lighting to reduce the effect of shadows between automobiles, rows of freight cars, piles of material, and similar bulky objects (see Table 26-25 for other recommended values).

The beam-spot areas at a (100-ft) distance *D* in formula A are given in Table 26-25 for various beam spreads and aiming angles of usual equipment having symmetrical candlepower distribution. In this table, *D* is the perpendicular distance measured from the floodlight to the plane of the lighted surface. *L* and *W* are the lengths and widths of the ellipses formed when floodlights are aimed at an angle to the lighted surface. At 0° the area is assumed to be circular; at other angles, it is elliptical. At other distances and spreads and for similar beam spreads and aiming angles, the spot areas vary as the square of the distance *D*, while *L* and *W* vary as the distance *D*. For example, if *D* is (24.38 m / 80 ft) and a 30 beam-spread floodlighting unit is aimed at a 50° angle, the elliptical spot areas computed from Table 26-22 will be $(80)^2/(100)^2 \times 9978$, or 6386 ft^2. Likewise, the length *L* of the ellipse will be $80/100 \times 144.4$, or 115.5 ft, and the width *W* will be $80/100 \times 87.96 = 70.4$ ft.

Formula B:

$$\text{Number of floodlights needed} = \frac{\text{illuminance} \times \text{area lighted}}{\text{light-loss factor} \times \text{utilization factor} \times \text{beam lumens}}$$

The light-loss factor allows for dust, dirt, and normal lamp depreciation. This is found under average conditions to be about 0.7. However, it may be as low as 0.3 for extremely dirty locations, where dust, dirt, and smoke are frequently suspended in the air.

The utilization factor, or coefficient of beam utilization, is the ratio of the lumens effectively lighting an area to the beam lumens, and it can be estimated from the following conditions:

1. If half or more than half of the floodlights are aimed so that all their beam lumens fall within an area, the overall utilization factor will be about 0.75.
2. If one-quarter to one-half of the floodlights are aimed so that all their beam lumens fall within an area, the overall utilization factor will be about 0.60.
3. If fewer than one-quarter of the floodlights can be aimed so that their beam lumens fall within an area, the overall utilization factor is likely to be not more than 0.40.

Most floodlights and projector- and reflector-type lamps as listed in the manufacturer's catalog are rated in beam lumens. These lumen ratings usually include only the light flux in that part of the

TABLE 26-25 Typical Recommended Illuminance Levels for Floodlighting, with Recommended Coverage Factors

Area	Footcandles	Lux	Minimum coverage factor
Building			3–4
General construction	10	100	
Excavation work	2	20	
Building exteriors and monuments:			2
Floodlighted			
Bright surroundings			
Light surfaces	5	50	
Medium-light surfaces	7	70	
Medium-dark surfaces	10	100	
Dark surfaces	15	150	
Dark surroundings			
Light surfaces	2	20	
Medium-light surfaces	3	30	
Medium-dark surfaces	4	40	
Dark surfaces	5	50	
Bulletin and poster boards:			1–2
Bright surroundings			
Light surfaces	50	500	
Dark surfaces	100	1000	
Dark surroundings			
Light surfaces	20	200	
Dark surfaces	50	500	
Open parking facilities*			
Basic	1	10	
Enhanced security	2.5	25	
Covered parking facilities			
General parking and pedestrian areas			
Day	5‡	50‡	
Night	5†	50†	
Ramps and corners			
Day	2	20	
Night	1	10	
Entrance areas			
Day	50‡	500‡	
Night	1	10	
Service station (at grade)			3–4
Dark surrounding			
Approach	1.5	15	
Driveway	1.5	15	
Pump—island area	5	50	
Building faces (exclusive of glass)	2	20	
Service areas	2	20	
Light surrounding			
Approach	2	20	
Driveway	2	20	
Pump—island area	10	100	
Building faces (exclusive of glass)	3	30	
Service areas	3	30	

*Based on recommended illuminance for pedestrian safety, Uniformity ratio 5:1.
†Minimum on pavement.
‡Minimum on pavement, sum of electric lighting and daylight.
§Vertical.

beam in which the candlepower values are 10% or more of the maximum candlepower of the floodlight.

As a general rule, it is wiser to design a system with a small number of floodlights with larger, more efficient lamps. This makes a simpler system to install, to control, and to maintain. Also from a control-of-light point of view it is desirable to choose a floodlighting unit having as narrow a beam spread as can be used and still maintain the coverage-factor requirements. It should be remembered, however, that large floodlights are hard to conceal; this is important where their daytime appearance may be objectionable architecturally.

To solve formula B, after choosing the desired illuminance and determining the light-loss and utilization factors, a size of floodlight is chosen for trial calculation and its beam lumens are substituted in the equation.

When the dimensions or shape of an area lead to the use of several types of floodlights, with different beam spreads, it is customary to divide the area into sections and plan a system for each of them. Buildings with setbacks are typical examples, also very tall structures such as towers or monuments. In setback buildings, you would design one setback at a time, selecting the type of floodlight most suitable for each. With towers or monuments, a similar approach is in order.

Floodlighting Design Problem. Assume that an area 200 by 200 ft is to be lighted to 5 fc. Also assume that the floodlights will be mounted on poles that are 60 ft high located on opposite sides of this area. By using a scale drawing to represent the luminaires placed 60 ft high and aimed toward the center line of the 200-ft-wide work space, the aiming angle will be about 60°. For trial-computation purposes it is best to start with the assumption that a narrow-beam-spread floodlighting unit will serve the coverage requirements and then change to a wider beam if found desirable. Hence, for this work space the lighting-design procedure is as follows:

1. From Table 26-26 the area which can be lighted for a 15° floodlight mounted 60 ft high at an aiming angle of 60° is

$$\frac{(60)^2}{(100)^2} \times 4720 = 1700 \text{ ft}^2$$

TABLE 26-26 Spot Areas for Narrow-, Medium-, and Broad-Beam Floodlights at a 100-ft Distance*

	15-deg beam, narrow			30-deg beam, medium			50-deg beam, broad		
Aiming angle	Spot area	*L*	*W*	Spot area	*L*	*W*	Spot area	*L*	*W*
0	545	26.34	26.34	2,250	53.58	53.58	6,830	93.26	93.26
10	570	27.16	26.70	2,370	55.38	54.49	7,220	96.81	95.00
15	606	28.25	27.30	2,518	57.70	55.56	7,760	101.54	97.30
20	657	29.89	28.00	2,757	61.27	57.29	8,600	108.75	100.7
25	735	32.18	29.1	3,102	66.28	59.58	9,880	119.18	105.5
30	846	35.31	30.5	3,603	73.21	62.67	11,770	134.06	111.8
35	1,000	39.57	32.3	4,333	82.78	66.64	14,710	155.58	120.4
40	1,230	45.42	34.6	5,420	96.18	71.75	19,500	187.66	132.3
45	1,583	53.59	37.6	7,129	115.46	78.62	27,900	238.35	149.0
50	2,115	65.34	41.3	9,978	144.43	87.96	44,810	326.58	174.7
55	3,043	82.97	46.7	15,160	190.84	101.14	87,140	509.39	217.8
60	4,720	111.10	54.1	25,880	273.21	120.6	265,100	1072.99	314.6
65	8,165	160.19	64.9	54,480	447.94	154.85			
70	16,800	258.97	82.6	180,910	1000.2	230.3			

*The spot area for any other distance can be computed by multiplying the area in this table at the selected aiming angle by D^2 and dividing by 10,000.

2. Solving formula A

$$\text{No. of 15° units needed for coverage} = \frac{(\text{area}) \times (\text{coverage factor})}{\text{spot area}} = \frac{40{,}000 \times 3}{1700} = 70$$

3. Solving formula B. A light-loss factor of 0.3 is found for this dirty location, and a utilization factor of 0.7 may be assumed for narrow-beam floodlights in this area. If a 1000-W incandescent unit is selected for a trial computation, and it has 9500 beam lm. Then the number of 1000-W floodlights needed =

$$\text{1000-W floodlights needed} = \frac{(\text{illuminance}) \times (\text{area})}{\text{LLF} \times \text{UF} \times \text{beam lumens}}$$

$$= \frac{5 \times 40{,}000}{0.3 \times 0.7 \times 9500} = 100$$

4. Since the number of floodlights (100) in formula B is greater than the number needed for adequate coverage, it could be concluded that, if 50 floodlight units are conveniently located on either side of the area, a satisfactory 5-fc lighting installation would be provided. On the other hand, if a minimum number of floodlights are desired, formula B can be resolved on the basis of using 1500-W luminaires having 12,300 beam lm. In this case,

$$\text{1500-W floodlights needed} = \frac{5 \times 40{,}000}{0.3 \times 0.7 \times 12{,}300} = 77$$

Hence, this work area can be satisfactorily lighted with 5 fc maintained in serviced if thirty-nine 1500-W 15° floodlights are well distributed along each side of the area at 60-ft mounting heights.

As seen from the foregoing example, it is possible to use a fewer number of higher-wattage luminaires when the number of units (formula B) to provide the required illuminance is considerably greater than the number required for adequate coverage (formula A). On the other hand, when the trial computations show that the number of luminaires to provide the required illuminance is less than those for adequate coverage, it becomes necessary to recalculate formula B by using beam lumens from smaller-sized units, until one is found which brings the answer equal to or greater than that for formula A. In other words, the answer to formula B should preferably never be less than that for formula A in order to provide adequate illumination as well as satisfactory coverage. In all cases, consideration should be given to the use of lamps with the highest luminous efficacy of a color suitable for the application if HID floodlights of appropriate beam spreads are available

A more detailed point-by-point floodlighting design could also be computed using software. In any floodlighting installation, care should be taken to avoid light trespass and light pollution (light directed to the sky). Local outdoor lighting ordinances may restrict illuminance levels or the use of particular types or mounting heights of outdoor lighting equipment.

26.15 ECONOMICS OF LIGHTING

The Value of Lighting. The value of good lighting depends on its use in the various fields of application.

Reports from a variety of industrial plants have indicated that better lighting created more satisfactory working conditions which in many instances meant an increase in production, reduced, spoilage, fewer accidents, and less labor turnover. Likewise, in stores and other selling areas, good lighting properly applied has been recognized as a necessity, not only to permit the customer to properly inspect the merchandise but also to direct his or her attention to displayed items and thereby increase the volume of sales. Moreover, light and lighting are an integral part of many modern buildings where the architect has incorporated luminous elements for their functional use and aesthetic value.

Reduced office productivity can be caused by poor lighting. In addition, workers using computers can experience headaches, eye strains, and other ergonomics-related problems from improper lighting.

For utilitarian installations, which have resulted in an increase in factory production or an increase in sales in a store, a dollars-and-cents value can readily be given to better lighting, but this is rarely something that can be predicted prior to system installation. On the other hand, there is wide acceptance of light and lighting for its humanitarian and decorative aspects, which are of inestimable value.

Cost of Lighting. The overall items for comparing different lighting systems must include both the initial and the operating cost. While one of these may be a dominant factor in the final selection, it is usually desirable to combine the two into some type of "total cost" indicator.

The computation of initial, operating, and total annual cost for various systems considered for a given interior must be based on certain common assumptions, if the systems are to be fairly compared. Some of the important considerations are:

1. Equal illuminance results—since different systems may not produce equal illumination levels in service, all costs should be equated to an equal maintained illuminance basis.
2. Equal rates for amortizing the initial investment and for addressing interest, taxes, and insurance should be used.
3. Equivalent electrical-energy rates and operating conditions (burning hours per year and starting frequency of the lamps) should be applied to different systems.
4. Cleaning schedule should be appropriate to each type of system.
5. Uniform labor rates among systems should be used for estimating the cost of installations, cleaning, and relamping. Some users request life-cycle costs.

Table 26-27 tabulates a cost-analysis procedure for comparing the cost of two or more lighting systems and is self-explanatory. Table 26-27 develops capital expense based on initial costs divided by the assumed year's life, and operating expenses computed from present energy, lamp replacement, cleaning, and repair costs. A simple payback analysis may also be computed from the differential first cost between two systems and the annual savings that a system provides. Life cycle costing is preferred by many owners, because it considers the present value of money, which depends on the rate of interest, and provides the most accurate measure of economic performance (see IESNA publication RP-31-96, *Economic Analysis of Lighting*).

Group-Relamping Costs. The cost of lamp replacement is made up of the cost of the lamp and the cost of labor required to replace the lamp. When the sum of the costs is reduced, of course, the total annual cost of operating the lighting system is reduced. It is difficult to assess the exact overall cost reduction without having all the other facts about the installation, but it should be kept in mind that economical lamp replacement means better overall lighting economics.

With spot replacement, the total replacement cost per lamp is equal to the cost of the lamp plus the labor cost of replacement. Group-relamping cost, to compare with this, is equal to the lamp cost plus group-relamping labor cost plus the cost of any interim spot replacements, divided by the group relamping interval to put both systems on an equal time basis.

REFERENCE ON ECONOMICS OF LIGHTING

1. *Economic Analysis of Lighting*, RP 31–96. IESNA 1996.

TABLE 26-27 Cost-Analysis Outline for Lighting Systems

General information	Lighting system 1	Lighting system 2
Installation data	…	…
Type of installation	…	…
Number of rows	…	…
Luminaires per row	…	…
Lamps per luminaire	…	…
Number of lamps	…	…
Watts per luminaire (including accessories)	…	…
Total watts	…	…
Maintained illuminance	…	…
Calculation of complete expense		
Capital expense	…	…
Estimated cost of each luminaire installed	…	…
Estimated wiring cost per luminaire	…	…
Cost per luminaire (luminaire plus wiring)	…	…
Number of luminaires	…	…
Total cost	…	…
Assumed years life	…	…
Total cost per year of life	…	…
Interest on investment (per year)	…	…
Taxes (per year)	…	…
Insurance (per year)	…	…
Total capital expense per year	…	…
Energy expense	…	…
Total watts	…	…
Average hours used per year	…	…
kWh per year	…	…
Average rate per kWh	…	…
Total energy expense per year	…	…
Lamp-renewal expense:	…	…
Number of lamps	…	…
Hours used per year	…	…
Total lamp hours per year	…	…
Rated lamp life, h	…	…
Average lamp renewals per year	…	…
Net price each	…	…
Replacement expense each (labor)	…	…
Net price plus replacement expense each	…	…
Total lamp-renewal expense per year	…	…
Lighting Calculation of complete expense	Lighting system 1	Lighting system 2
Cleaning expense	…	…
Number of washings per year	…	…
Worker-hours each (est.)	…	…
Worker-hours for washing	…	…
Number of dustings per year	…	…
Worker-hours each (est.)	…	…
Worker-hours of dusting	…	…
Total worker-hours	…	…

TABLE 26-27 Cost-Analysis Outline for Lighting Systems (*Continued*)

Lighting Calculation of complete expense	Lighting system 1	Lighting system 2
Expense per worker-hour	…	…
Total cleaning expense per year	…	…
Repair expense	…	…
Repairs (based on experience, allocation of repair worker's time, etc.)	…	…
Estimate of total repair expense per year	…	…
Lamp disposal expense per year*	…	…
General information		
Recapitulation	…	…
Total capital expense per year	…	…
Total energy expense per year	…	…
Total lamp renewal expense per year	…	…
Total cleaning expense per year	…	…
Estimate of total repair expense per year	…	…
Estimate of total lamp disposal expense per year	…	…
Complete lighting expense for year	…	…

*An essential ingredient of fluorescent and HID lamps is a small amount of mercury. More strict regulations now apply to the disposal of such lamps, to minimize the possibility of contaminating air and water.

26.16 LIGHTING MAINTENANCE

Maintenance of Lighting. Good lighting system maintenance is good economics, that is, good maintenance assures users that they actually get the light they pay for. With a well-maintained lighting system, the user also gets the better conditions that the system was designed to provide. These better seeing conditions contribute directly to higher productivity and improved morale in factories, offices, and schools. In stores, the better lighting that results from good maintenance helps to increase sales. In addition, all areas benefit from a better appearance and fewer work interruptions that come with good maintenance.

Causes of Light Loss. Several factors contribute to light loss, and the effects of these factors vary with the kind of activity that takes place and the location of the establishment. For example, spaces vary as to the amount and type of dirt present in the air. Obviously, the amount of dirt in a machining factory is greater than that in an air-conditioned office. It is important to recognize these variations in considering the light losses which result from dirt on lamps and lighting units and dirt on room surfaces. These two, along with the unavoidable lamp-lumen depreciation, are the principal factors that cause light loss in every lighting installation.

Benefits of Cleaning. A lighting system should not only be cleaned properly, it should be cleaned at the proper time. This combination produces a cleaning program, which is a profitable investment because three principal benefits can be obtained.

1. *Better energy utilization for lighting*. Dirt absorbs light—cleaning removes the dirt and thus helps to maintain the light level. As a result, visibility is maintained, which benefits the user.
2. *Reduced maintenance costs*. A good maintenance program calls for cleaning the lighting system at the most economical time. It makes use of the most efficient methods and equipment. In this way the time and materials required are reduced, thus maintenance costs are lowered.

3. *Better appearance*. Clean lighting systems improve the appearance of the space. This is conducive to improved morale, better housekeeping, and increased occupant satisfaction.

To further improve the appearance of the area being lighted, walls and ceilings should be periodically cleaned and repainted as part of the cleaning program.

Relamping Benefits. The lamps in a lighting system can be replaced individually as they burn out, or the entire installation can be replaced before the lamps reach their average life. Individual replacement is usually called *spot replacement*; mass replacement is called *group relamping*. The labor costs saved by group relamping in large installations and in many small ones more than compensate for the value of the depreciated lamps that are thrown away before they burn out. Other advantages which always accompany group relamping are more light, fewer work interruptions, better appearance of the lighting system, and less maintenance of auxiliary equipment.

There are five principal advantages for group relamping. The first three apply to all lighting systems, the last two chiefly to fluorescent and HID systems.

1. *Reduced labor costs often mean net savings.* Group relamping saves on labor costs, largely because much of the travel and setup time required to change lamps individually is eliminated.
2. *More light delivered*. All lamps depreciate in lumens continually as they burn. The earlier they are replaced, the higher the maintained illumination will be without adding to the use and cost of electric energy since a larger number of luminaires are needed to achieve a target design illuminance.
3. *Fewer work interruptions*. Group relamping can be done at a convenient time—during vacation shutdowns or after working hours, for example, when there will be no interruption of operations. The number of interruptions to report burnouts or to replace tem is greatly reduced.
4. *Better appearance of the lighting system*. Black ends, color variations, and differences in brightness between adjacent old and new fluorescent lamps are common when spot replacement is used. With group relamping, all the lamps are the same age, and appearance is more uniform.
5. *Less maintenance of auxiliary equipment*. Abnormal operating conditions that may occur at the end of lamp life can damage starters and ballasts. When most of the lamps are replaced before they reach the end of life, auxiliary equipment lasts longer.

26.17 LIGHTING MEASUREMENT DEVICES

A variety of light measurement devices are available to a lighting practitioner for use in measuring lighting system performance. The most common devices are meters to measure illuminance and luminance.

Illuminance Meter. An illuminance meter is generally constructed with a silicon or selenium cell placed beneath an integrating white diffuser that may be either flat or round. The cells generate a potential difference when irradiated which bears a relationship to the level of illumination. The cells also have a spectral sensitivity that is chiefly in the visible and near ultraviolet region. Because it is different than that of the human eye, to measure photometric quantities (quantities based on the definition of the lumen) the cell must contain a color-correcting filter to align the detector's performance with that of the CIE luminous efficiency curve (see Fig. 26-38). Illuminance meters must also provide a reading that follows the cosine law of illuminance. That is, that the illuminance is proportional to the cosine of the angle the incident light ray makes with the normal to the measurement surface. Most illuminance meters perform well in this regard, but are less accurate at very high incident light angles (above 75° to 80°). When using an illuminance meter, it is important to properly orient the meter and to avoid shadowing the light arriving at the meter.

Luminance Meter. A luminance meter is used to measure the light leaving a surface in a particular direction. This type of meter must also be color corrected, and is used to record the luminous intensity (candelas) per projected area leaving the surface in the direction of the meter. Luminance meters are typically configured with a lensed optical system and an eyepiece through which the operator views the surface. A circle in the center of the image viewed through the meter indicates the surface area to be measured. The angular extents of this measurement zone may vary, but it is typically on the order of 1° or some fraction of a degree. When using a luminance meter, it is important that the operator focus the image using an adjustable outer lens. Since it may be difficult to hold such a meter steady and achieve a static reading, these meters typically have a trigger that may be pulled to make a reading, then released to freeze the reading.

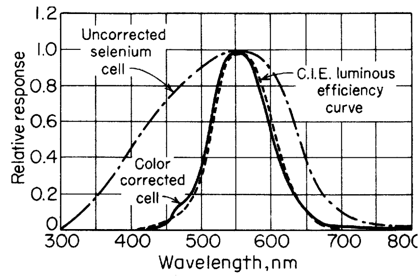


FIGURE 26-38 Average spectral sensitivity characteristics of selenium barrier-layer cells, compared with CIE spectral luminous efficiency curve.

Other more specialized devices for measuring lighting quantities are also available. These are unlikely to be needed by a lighting designer, but may be useful to manufacturers, research laboratories, and other lighting specialists. This list of devices include chromaticity meters, exposure meters, reflectometers, goniophotometers, glossmeters, fluorometers, colorimeters, refractometers, and pyranometers.

References on illumination may be found in texts on the subject of illuminating engineering and in appropriate journals. A few are listed in the bibliography.

BIBLIOGRAPHY

ASHRAE/IESNA Standard 90.1/2004: Energy Standard for Buildings Except Low-Rise Residential Building, American Society of Heating, Refrigerating and Air-Conditioning Engineers & Illuminating Engineering Society of North America, 2004.*

Ander, G.D., *Daylighting Performance and Design*, 2d ed., 2003.*

Boyce, P., *Human Factors in Lighting*, 2d ed., 2003.*

Cuttle, C., *Lighting by Design*, 2002.*

Deveau, R.L., *Fiber Optic Lighting: A Guide for Specifiers*, 2000.*

Egan, D., and V., Olgvay, *Architectural Lighting*, 2d ed. 2002.*

Eley, C., Technical Ed., *Advanced Lighting Guidelines*, New Buildings Institute, 2003.*

Fetters, J.L., *Handbook of Lighting Surveys and Audits*, 1997.*

Gordon, G., *Interior Lighting for Designer*, 4th ed., 2003.*

IES Lighting Handbook, IES Recommended Practices and IES Design Guides (separate publications in many fields), IES Lighting Education materials, Illuminating Engineering Society of North America.

Karlen, M., and J., Benya, *Lighting Design Basics*, 2004.*

Kay, G., *Fiber Optics in Architectural Lighting: Methods, Designs, and Applications*, 1999.*

Leslie, R.P., and P.A., Rodgers, *Outdoor Lighting Pattern Book*, 1997.*

Lighting Research Center, Rensselaer Polytechnic Institute, Series of publications identified as guides, case studies, technical reports, lighting futures newsletters, national lighting product information, lighting answers, delta brochures, etc. www.lrc.rpi.edu

Michel, L., Light: *The Shape of Space—Designing with Space and Light*, 1995.*

Moyer, J.L., *The Landscape Lighting Book*, 2d ed., 2005.*

Murdoch, J., *Illumination Engineering: From Edison's Lamp to the LED*, 2003.*

Steffy, G., and G., Woodall, *Time-Saver Standards for Architectural Lighting*, 2000.*

Steffy, G., *Architectural Lighting Design*, 2d ed., 2002.*

Wood, D., *Lighting Upgrades: A Guide for Facility Managers*, 2d ed., 2004.*

Wyszecki, G., and Stiles, W.S., *Color Science*, Wiley, 1982.

Journals

Lighting Design and Application

LEUKOS, *The Journal of the Illuminating engineering Society of North America.*[†]

Lighting Research & Technology[†]

Journal of the Opical, Society of America.

---

* May be obtained from the Illuminating Engineering Society of North America, 120 Wall Street, 17th Floor, New York, NY 10005-4001.

[†]Formerly, Illuminating Engineering, Transactions of the Illuminating Engineering Society, and the *Journal of the Illuminating Engineering Society*.

SECTION 27
LIGHTNING AND OVERVOLTAGE PROTECTION

A. P. (Sakis) Meliopoulos
Professor, School of Electrical and Computer Engineering, Georgia Institute of Technology

CONTENTS

27.1 INTRODUCTION

Temporary overvoltages in power systems occur for a variety of reasons such as faults, switching, and lightning. By far, the most severe overvoltages result from lightning strokes to the power system. Most likely, lightning overvoltages will be very high, resulting in insulation breakdown of power apparatus with destructive results. It is therefore imperative that power systems be designed in such a way that expected overvoltages be below the withstand capability of power apparatus insulation. Many times, this basic requirement is translated into excessive cost. For this reason, one seeks a compromise in which power systems are designed in such a way that the possibility of destructive failure of power apparatus due to overvoltages is minimized. This procedure is based on coordinating the expected overvoltages and the withstand capability of power apparatus. Two steps are typically involved: (1) proper design of the power system to control and minimize the possible overvoltages and (2) application of overvoltage protective devices. Collectively, the two steps are called *overvoltage protection* or *insulation coordination*.

The importance of overvoltage protection cannot be emphasized enough. First it affects system reliability, which translates into economics. Traditionally, overvoltage protection methods were guided by the objective to maximize system reliability with reasonable investment cost. In this sense, transient overvoltages which do not lead to interruptions are acceptable and short-duration interruptions are tolerable. Recently, however, with the introduction of sensitive electronic equipment, new concerns have been raised. The issue of power quality is important and it is transforming the practices for overvoltage protection. While the application of overvoltage protection devices is pertinent, more and more emphasis is placed on design procedures to minimize the possible overvoltages and control the sources of disturbances. An attempt has been made in this section to provide a balanced treatment of overvoltage protection in view of present-day concerns.

The subject of lightning and overvoltage protection is rather complex. A thorough treatment requires good understanding of many related subjects. First, the mechanisms by which lightning is generated and how its pertinent characteristics are related to power systems must be well understood. Second, the response of power systems to lightning and other causes of overvoltages must be studied. Analysis methods to study the phenomena are indispensable tools, which provide the basis for proper selection of design options. Invariably, overvoltages can be minimized, but they cannot be eliminated. As a result, power systems must be protected against overvoltages using overvoltage protection devices (surge arresters). In recent years, major breakthroughs have occurred in protective device technology. Effective protection requires a deep understanding of the capabilities of present technology as well as its limitations.

27.2 BASIC CONCEPTS AND DEFINITIONS

Electric power systems are subjected to external surges (lightning) as well as internally generated surges (switching), which may result in temporary high voltages. To maintain a highly reliable system, protection against these overvoltages is needed. This need is dictated by the fact that the insulation of power equipment (which may be air, oil, SF_6, etc.) is subjected to breakdown if sufficiently high voltage is applied. This protection involves a coordinated design of the power system itself and placement of proper protection devices at strategic locations for the purpose of suppressing overvoltages and avoiding or minimizing insulation failures.

Coordinated design involves

Effective grounding techniques

Use of shielding conductors

Preinsertion resistors during switching

Switching angle control among breaker poles

Use of surge capacitors

Protection devices include spark gaps and various designs of surge arresters.

The basic objective of overvoltage protection of power systems is to avoid insulation breakdown and associated outages or damage to equipment. The most common insulators used in power system apparatus and their characteristics are listed in Table 27-1.

In general, in terms of potential damage to equipment, the insulation of power apparatus can be classified into external and internal as follows:

- External insulation

  Air

  Porcelain

  Glass

- Internal insulation

  Oil

  SF_6

  Mica

The effects of external insulation breakdown are not as destructive as internal insulation breakdown. The reason is that external insulation is, in general, self-healing (self-restoring) after the cause of breakdown (overvoltage) ceases to exist. On the other hand, internal insulation breakdown generally results in permanent damage to the equipment and possibly catastrophic failure. These facts dictate different approaches for external and internal insulation protection. For external insulation protection, the objective is to minimize the expected number of insulation breakdowns subject to economic constraints. In this sense, many sophisticated approaches have been developed, which

TABLE 27-1 Common Insulators in Power Apparatus

Insulator	Breakdown, MV/m	Resistivity, $\Omega \cdot$ m	Relative permitivity
Air	3	$\rho = \infty$	$\epsilon_r = 1$
Oil	10	$10^4 \times 10^{10}$	2.2
SF_6	15 at 1 atm 59 at 5 atm		
Mica	100	$10^{11} - 10^{15}$	4.5–7.5
Porcelain	10	3×10^{12}	5.7
Glass	…	10^{12}	4–7

balance system reliability (which is mainly related to insulation breakdowns) versus cost. Because many of the exogenous parameters, such as lightning strength and soil parameters are statistical in nature, the methodologies use statistical approaches. For internal insulation protection, deterministic methods are applied where the objective is to design for zero insulation breakdowns.

The above simplistic characterization of external and internal insulation is not always apparent in power apparatus. Specifically, the insulation of a specific power apparatus may be complex. For example, consider a transformer. The windings of the transformer may be submerged in oil (the dielectric is oil) while the terminals are exposed to air through the bushings (the dielectric is the air). When considering withstand capability of a power apparatus, we are not concerned with which dielectric will break first, although this is part of the design process. But rather we are concerned with the question of at what voltage the insulation (any part) will break down. Because insulation breakdown depends on voltage waveform as well as on some other factors, the following definitions, which have been taken from the ANSI Std C92.1, apply:

Withstand voltage. The voltage that electrical equipment is capable of withstanding without failure or disruptive discharge when tested under specified conditions.

Insulation level. An insulation strength expressed in terms of a withstand voltage (typically 10% less than the withstand voltage).

Transient insulation level (*TIL*). An insulation level expressed in terms of the crest value of the withstand voltage for a specified transient wave shape, for example, lightning or a switching impulse.

Lightning impulse insulation level. An insulation level expressed in terms of the crest value of a lightning impulse withstand voltage.

Switching impulse insulation level. An insulation level expressed in terms of the crest value of a switching impulse withstand voltage.

Basic lightning impulse insulation level (*BIL*). A specific insulation level expressed in terms of the crest value of a standard lightning impulse.

Basic switching impulse insulation level (*BSL*). A specific insulation level expressed in terms of the crest value of a standard switching impulse.

Note that two of the most commonly used measures, the basic lightning impulse insulation level and the basic switching impulse insulation level, are the most widely used values to characterize the insulation of power apparatus. Note that they are defined in terms of two specific waveforms: (1) the standard lightning impulse and (2) the standard switching impulse. The definitions of these waveforms are

Standard lightning impulse. A full impulse having a front time of 1.2 μs and a time to half value of 50 μs. It is described as a 1.2/50 impulse. (See *American National Standard Measurement of Voltage in Dielectric Tests,* C68. 1.)

Standard switching impulse. A full impulse having a front time of 250 μs and a time to half value of 2500 μs. It is described as a 250/2500 impulse. (See American National Standard C68.1.)

These waveforms are illustrated in Fig. 27-1.

The standard impulses were introduced because they remotely resemble lightning and switching waveforms, and they can be easily generated in a laboratory via an impulse generator. The basic structure of an impulse generator is illustrated in Fig. 27-2*a*. By stacking many basic structures together, one can create an impulse generator capable of generating an output impulse many million volts in crest.

The impulse voltage withstand of a power apparatus is strongly dependent on the duration of the impulse voltage. The time dependence is mainly due to the fact that arc generation involves an electron avalanche which takes a finite time to form. The full development of an arc across an insulator is classified as a breakdown. The time to breakdown is normally quantified with a volt-time characteristic. This characteristic can be determined by applying impulses across an insulator of increasing magnitude and recording the voltage and time at which breakdown occurred. For self-restoring insulation, this test is relatively simple and is illustrated in Fig. 27-3. In this way, volt-time curves for all insulators in usage have been determined. Unfortunately, the withstand voltage of non-self-restoring insulation cannot be readily determined without destroying the sample. This means that determining

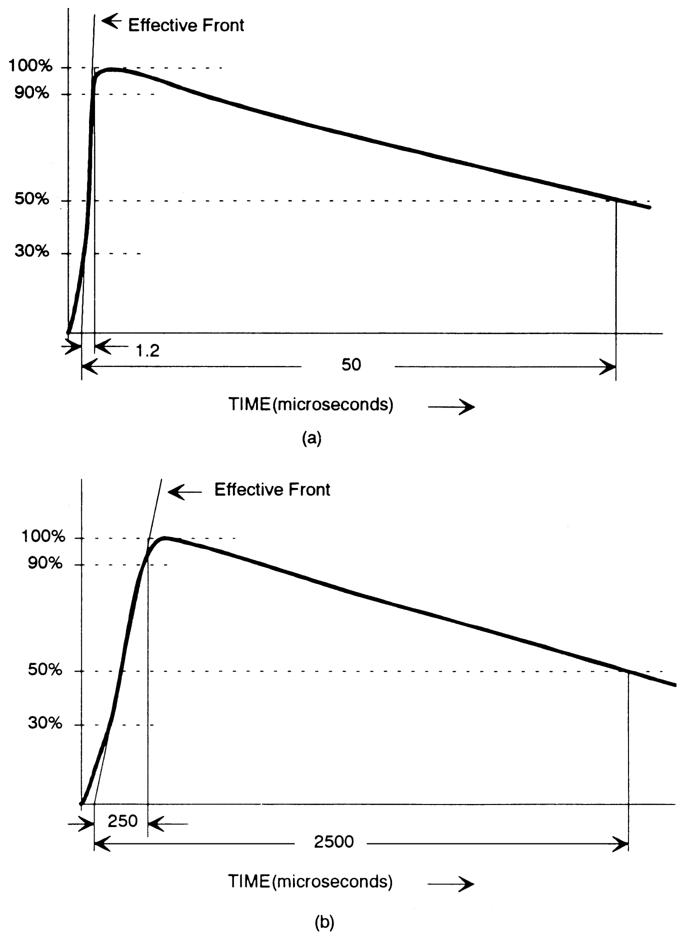


FIGURE 27-1 Standard waveform: (*a*) standard lightning impulse; (*b*) standard switching impulse.

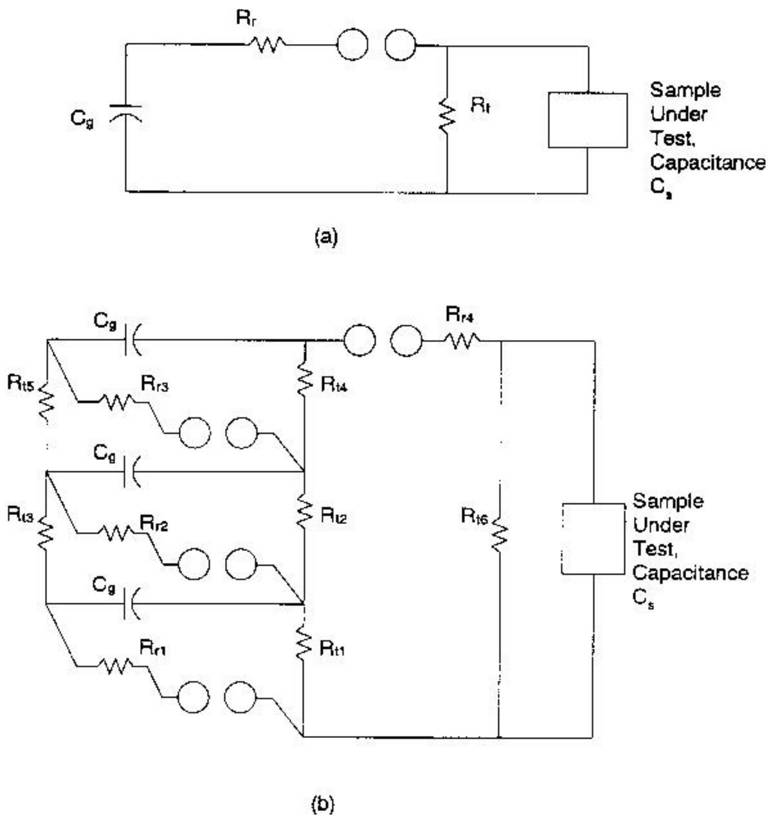


FIGURE 27-2 Impulse generator: (*a*) single-stage; (*b*) multiple-stage.

the volt-time curve of non-self-restoring insulation is a practical impossibility. For this reason, the methods for determining withstand voltage for internal insulation are different. Specifically, internal insulation is designed for a specific withstand capability, the *design withstand*. The manufacturer must guarantee a certain withstand at which the insulation, if tested, will not fail. This is the *tested withstand* and it is normally lower than the design withstand. Apparently, the *actual withstand* cannot be known without destroying the sample. The actual withstand is definitely higher than the tested withstand and probably higher than the design withstand.

There is another issue related to the fact that the withstand voltage depends on many other factors that exhibit random variations. Some of them are

Insulation geometry and smoothness of surfaces

Insulation contamination

Atmospheric conditions

Voltage polarity

It is a practical impossibility to quantify the effects of all variables on voltage withstand. For this reason, voltage withstand is described in statistical terms. In this sense, the following definitions apply with reference to Fig. 27-3:

Critical flashover (CFO) is the crest voltage of an applied impulse wave that will cause flashover on the tail of the wave 50% of the time and no flashover the other 50% of the time.

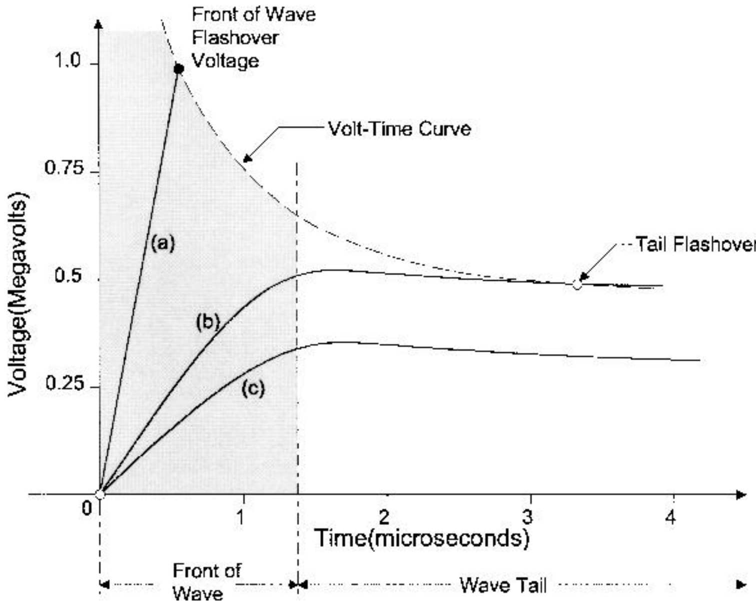


FIGURE 27-3 Determination of the volt-time curve of insulation breakdown.

Critical withstand is the highest crest voltage insulation can take without flashover under specified conditions—usually less than 1% probability of flashover.

Rated withstand is the crest voltage that insulation is required to withstand without flashover when tested by established standards under specified conditions (usually 5% to 10% less than critical withstand).

In summary, in this section we reviewed several basic concepts and definitions, which are useful in the process of designing protection systems.

27.3 MECHANISMS AND CHARACTERISTICS OF LIGHTNING

Introduction. Atmospheric electrical discharges known as *lightning* or *thunderbolts* (from cloud to cloud or cloud to ground) have captured the imagination and fear of the human race since ancient times. The ancient Greeks believed that lightning was Zeus' tool to punish human misbehavior or to demonstrate his anger. It was not until Benjamin Franklin that the first scientific inquiry occurred into the phenomenon of lightning. Since that time, lightning has been extensively studied and many theories have been developed, which reasonably explain the phenomenon. In addition to these theories, there exists an enormous amount of measured data of lightning characteristics. These data are useful for design of protection schemes against lightning.

This section presents a brief overview of the theory of thundercloud formation and lightning, the characteristics of lightning, and describes existing relevant data.

The Electrification of Thunderclouds. The cause of lightning is separation and accumulation of electrical charges in clouds via certain microphysical and macrophysical phenomena. This electrification results in electric field intensities high enough to cause air breakdown and subsequent development of lightning. To explain these phenomena, certain theories have been developed. The most useful are the precipitation and convection theories and later improvements, most notably the charge-reversal temperature theory. Understanding of these theories is helpful in the design of protection systems against lightning. A brief description of the cloud electrification theories is provided in this section.

The precipitation theory, postulated as early as 1885 by physicists Elster and Geitel, is based on the observation that large water droplets accelerate toward ground because of gravity, while smaller water droplets (mist) remain suspended in air or rise as warmer air moves upward. Collisions between large water droplets and mist of water droplets and possibly ice crystals in the colder altitudes result in transfer of a net negative charge to the large water droplets. As they move toward lower altitudes (by gravity), they cause a net negative charge in the lower part of the cloud. Conservation of charge requires that the upper part of the cloud be positively charged, resulting in a dipole structure in the thundercloud. A simplified illustration of the process is given in Fig. 27-4.

The convection theory, which was formulated much later, is based on transfer of charged particles from one location of a cloud to another by the upward and downward drafts in the cloud. The theory suggests that the charged particles are generated by two mechanisms: (1) cosmic rays impinge on air molecules and ionize them, resulting in two ions, one positively charged, the other negatively charged; and (2) high-intensity electric fields around sharp objects on the earth's surface produce corona discharges, which result in positively charged ions. The positive ions are transported to higher altitudes by the upward draft in the cloud. On the other hand, the negative ions attach themselves to water droplets and ice particles, which move to lower altitudes due to gravity or downward drafts. The net result is a dipole structure in the thundercloud. A simplified illustration of the process is given in Fig. 27-5.

Precipitation and convection occur in a thundercloud simultaneously. Yet the two theories are distinct and independent. Both theories postulate that the thundercloud is a dipole with the negative pole near the earth, that is, negative dipole. Measurements made by Wilson and later by Simpson of the polarity of the dipole resulted in conflicting conclusions which generated debate and further research. Specifically, Wilson's measurements indicate that the thundercloud is a negative dipole (negative charge at the lower part of the cloud) while Simpson's measurements indicated a positive dipole. It took five decades of additional experimentation and measurements to resolve this apparent conflict. Today's most complete theory for lightning phenomena has established the fact that the structure of a thundercloud is *tripolar*, not bipolar. This structure allows the understanding of both Wilson's and Simpson's conclusions. Specifically, an electric tripole of the size of a thundercloud observed from a single specific point will appear as a dipole. Depending on the point of observation

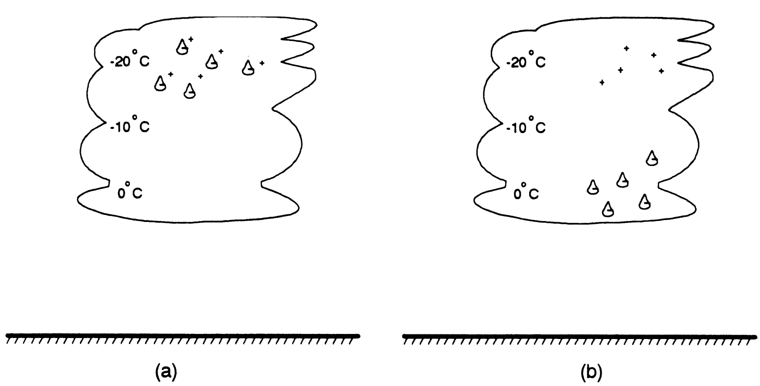


FIGURE 27-4 Illustration of the precipitation theory of cloud electrification: (*a*) separation of the charge due to collisions; (*b*) cloud electrification due to precipitation of charged water droplets.

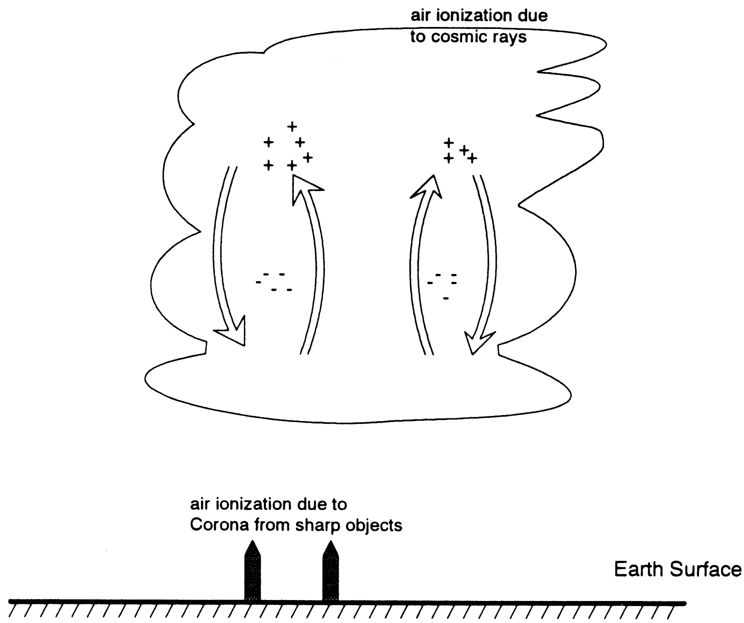


FIGURE 27-5 Illustration of the convection theory of cloud electrification.

one may conclude that it is a negative or positive dipole. Apparently, Wilson and Simpson made their measurements from different observation points. Their measurements were correct but because Wilson made the measurements from a distant point he concluded that the thundercloud is a negative dipole, while Simpson made his measurements from a point underneath the head of the cloud and concluded that the thundercloud is a positive dipole.

Both theories, precipitation and convection, do not completely explain all phenomena occurring in a thundercloud. For example, it has been observed from studies that larger droplets, when they break, acquire positive charge on aggregate. This leads to the hypothesis that the positive charge is due to large droplets that break as they accelerate toward ground. However, this hypothesis is not totally true because it does not explain the fact that precipitation particles below the negative charge carry much greater positive charge than those produced by the droplet fragmentation process. Another hypothesis was based on ice particles accelerating toward ground—as the ice particles reach lower altitudes, they melt and tend to acquire positive charges, which explains the existence of positive charge at altitudes below 4 km. However, this hypothesis still does not explain the existence of positive charges at higher altitudes.

Recent measurements and observations in the past three decades resulted in another hypothesis which explains the tripole nature of a thundercloud. This is the so-called charge-reversal hypothesis, which states that when graupel particles collide with ice crystals, the charge transferred to a graupel particle is dependent on the temperature. At temperatures above a certain value, which is called the *charge-reversal temperature*, the transferred charge is positive. The exact value of the charge-reversal temperature is being debated, but it is believed to be around −15°C. The process is illustrated in Fig. 27-6 in a simplified manner. Considering the fact that the temperature of the atmosphere is −15°C at an approximate altitude of 6 km, this means that due to collisions of graupel particles and ice crystals, the thundercloud will be, on aggregate, negatively charged for altitudes above 6 km and positively charged below 6 km. The situation is illustrated in Fig. 27-7. This hypothesis has been verified in the laboratory and explains the levels of negative and positive charges in a thundercloud. Yet, the exact microphysics of this phenomenon are practically unknown.

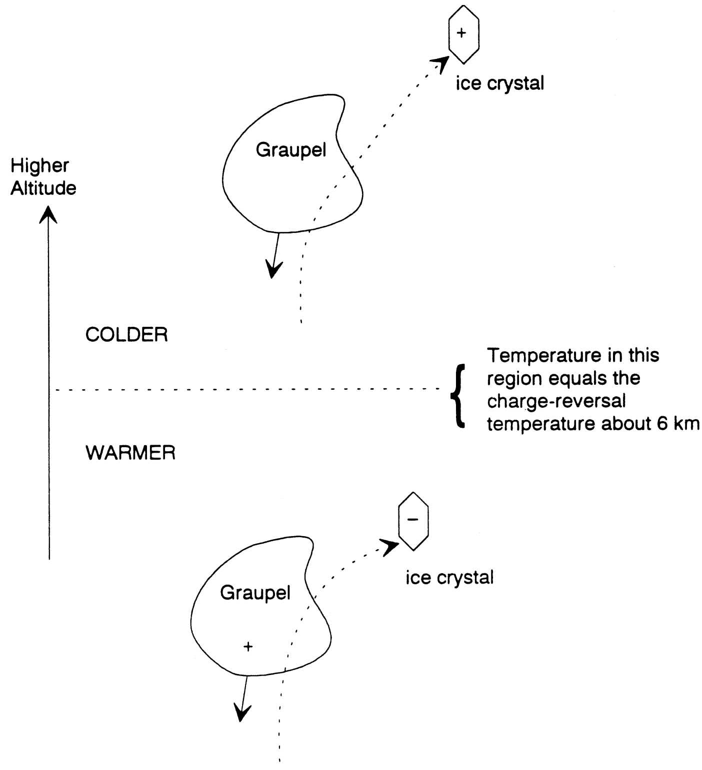


FIGURE 27-6 Explanation of the charge-reversal temperature theory.

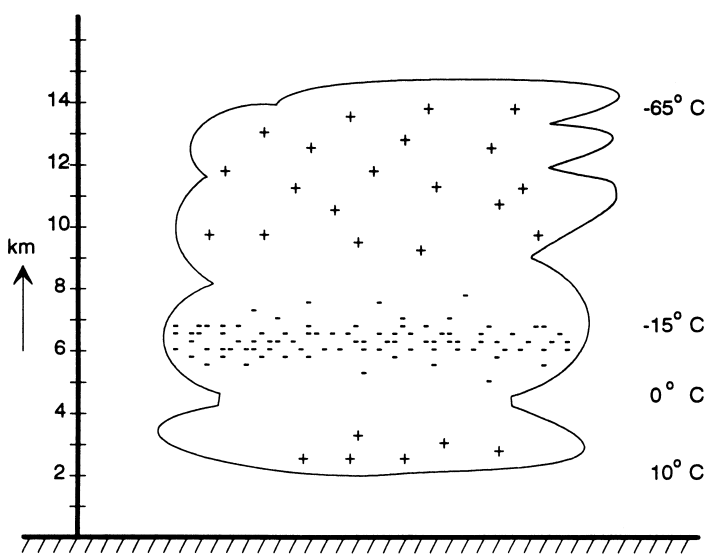


FIGURE 27-7 An electrified thundercloud is typically tripolar.

In summary, the precipitation model with the graupel—ice crystal interaction and charge-reversal temperature best explains most of the behavior of a thundercloud. Yet this model totally ignores the forceful upward and downward drafts within a thundercloud. The convection model considers these drafts but it is unable to explain certain observed phenomena in a thundercloud. Perhaps one day a theory will be developed, which combines the precipitation and convection models and completely accounts for all phenomena related to the electrification of a thundercloud. What has been verified with measurements are the following facts: a thundercloud can be electrified in such a way that positive charge accumulates at the top of the cloud and negative, at the lower part of cloud. A smaller positive charge may be present at lower altitudes of a thundercloud. These charges are responsible for lightning. The mechanism of lightning is explained next.

Mechanisms of Lightning. Lightning initiates whenever the charge accumulation in a thundercloud is such that the electric field between charge centers inside the cloud or between cloud and earth is very high. For power engineering purposes, only cloud-to-earth lightning strokes (ground flashes) are of importance and will be discussed next. An electrified thundercloud will generate an electric field in the space between the cloud and earth as is illustrated in Fig. 27-8. When the intensity of this field is high enough, a discharge will initiate. Typically, the process involves three phases. In the first phase, the high electric field intensity may generate local ionization and electric discharges, which are known as *pilot streamers*. A pilot streamer is followed by the so-called *stepped leader*. The stepped leader is a sequence of electric discharges, which are luminous; they propagate with a speed approximately 15% to 20% of the speed of light, and they are discrete, progressing approximately 50 m at a time. The time between *steps* is few microseconds to several tens of

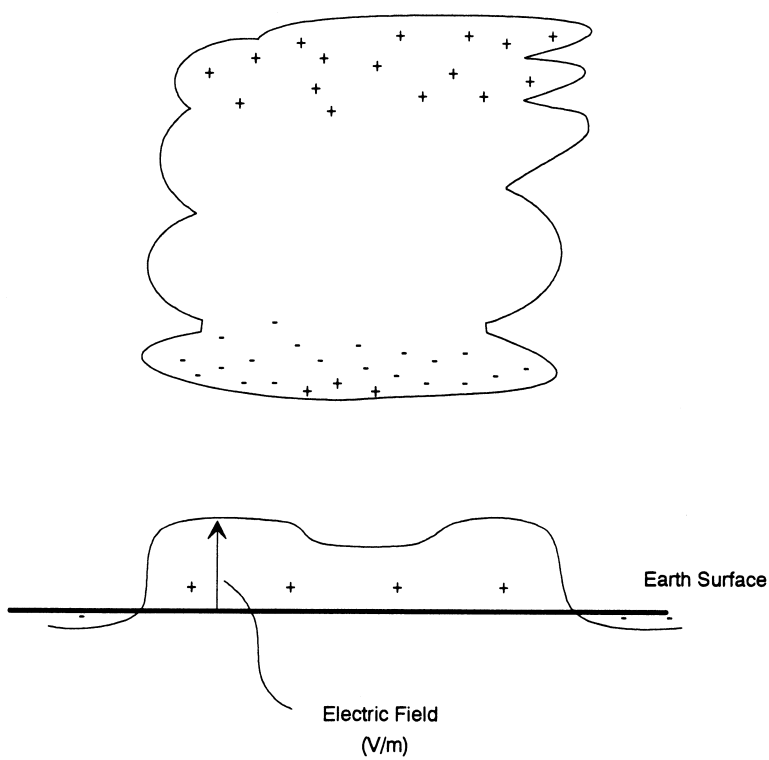


FIGURE 27-8 Illustration of electric field below an electrified thundercloud.

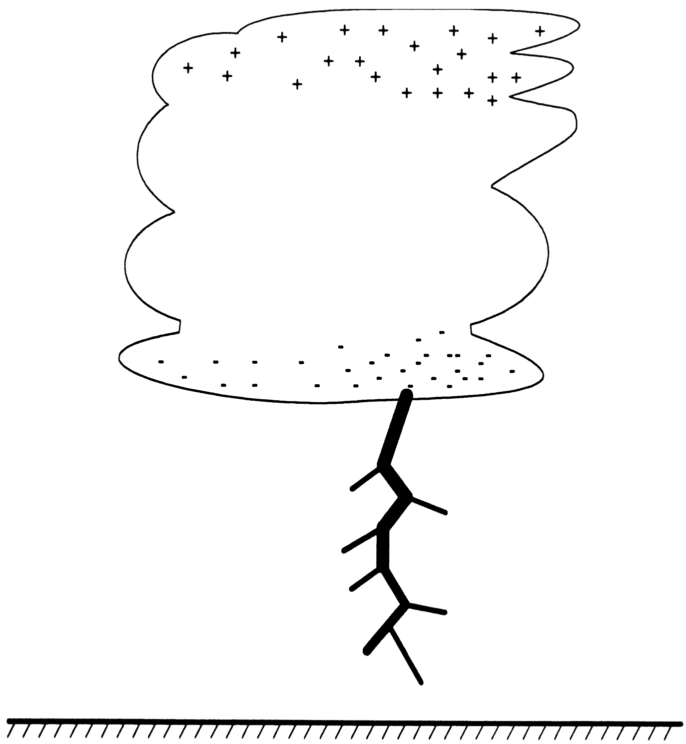


FIGURE 27-9 Illustration of stepped-leader development.

microseconds. A pictorial view of the stepped leader development is shown in Fig. 27-9. The stepped leader will eventually reach the surface of the earth and will strike an object on the earth. However, where it will strike is not determined until the stepped leader is within a *striking distance* from the object. A model for the striking distance will be described in Sec. 27.7. It is possible that a stepped leader or multiple stepped leaders may also initiate from an object on the surface of the earth. In this case, the two stepped leaders may meet at some point. There is also evidence that the initial stepped leader may originate from a tall structure on the earth and not from the cloud.

The second phase initiates when the stepped leader reaches an object on the earth or meets an upward moving stepped leader. Specifically, a high-intensity discharge occurs through the channel established by the stepped leader. This discharge is extremely luminous and therefore visible. It propagates with a speed of about 10% to 50% of the speed of light. The development of the return stroke is illustrated in Fig. 27-10. The return stroke carries an electric current of anywhere from few thousands of amperes to 200 thousands of amperes. The current magnitude rises fast, within 1 to 10 μs, to the peak value and then decreases rapidly. The discharge is known as the *return stroke* or simply the *lightning stroke*. The return stroke transfers a substantial amount of positive charge from the earth to the cloud and specifically to the charge center where the lightning was originated. This transfer results in a significant lowering of the potential of the charge center. This phenomenon initiates the third phase of lightning. In this phase, discharges may occur from other charge centers within the thundercloud to the depleted charge center because of the increased potential difference between them. This discharge will trigger another stroke between cloud and ground through the already established conductive channel with the first stroke. This process may be repeated several times, depending on the electrification status of the thundercloud, resulting in multiple strokes. There is evidence

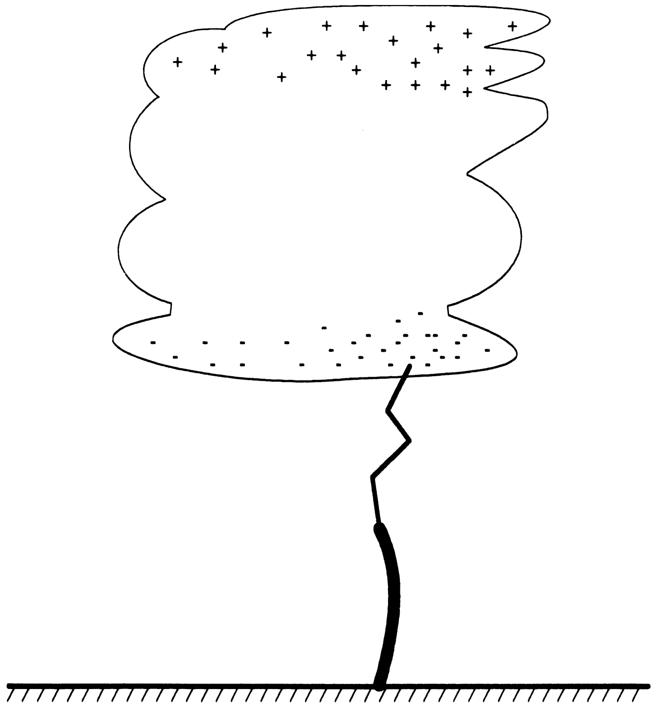


FIGURE 27-10 Illustration of return-stroke development.

that most lightning to ground involves multiple strokes. For example, an analysis of 1430 strokes to the earth by Anderson (1968) resulted in the following statistics:

Single-stroke lightning	36%
Lightning involving six or more strokes	21%
Mean value of multiple strokes	3

It should be mentioned, however, that positive polarity lightning is typically single stroke. Extreme cases have been recorded with a large number of multiple strokes such as 40 or 50 with duration of the entire lightning event approaching 1 s. For example, a 40-stroke lightning event, which lasted 0.624 s is on record. The time interval between successive strokes may be in the order of few milliseconds. However, there is evidence that sometimes the multiple strokes may be so smooth as to appear as a continuous lightning current. This occurs because as long as there is a conductive path between the cloud and ground, electric current will flow until the cloud is neutralized enough for the conductive path to interrupt. The continuous flow of lightning current can be destructive because of its long duration even if the magnitude may be much lower than the crest of a stroke.

Characteristics of Lightning Strokes. The parameters of lightning ground strokes are very important in the design of protection schemes against lightning. The most important parameters are

- Voltage
- Electric current
- Waveform
- Frequency of occurrence

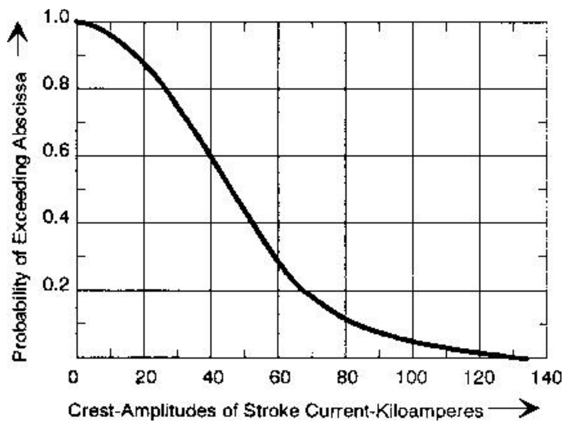


FIGURE 27-11 Distribution of lightning current magnitudes [recorded by Berger (1967) at Mount San Salvatore].

The voltage between a thundercloud and the earth prior to a ground stroke has been estimated from 10 to 1000 MV. For design work, however, the protection engineer is interested in the voltage appearing on the stricken power apparatus. This voltage will be equal to the product of the impedance times the stroke current.

It is generally accepted that the ground-stroke current is independent from the terminating impedance. The reason is that the terminating impedance is much lower than the resistance of the lightning discharge channel, which is on the order of few thousand ohms. Thus, a ground stroke is normally considered as an ideal current source at the point of strike. The crest of the stroke electric current can vary over a wide range: 1 to 200 kA. Data on ground-stroke current magnitudes have been collected by many researchers. Among those, the work of Berger (1967) at Mount San Salvatore in Switzerland has been widely accepted. Statistical representation of these data is shown in Fig. 27-11.

The waveform of the lightning ground stroke current, and especially the rise time, is very important. Again, statistical representation of stroke current rise times data collected by Berger is given in Fig. 27-12.

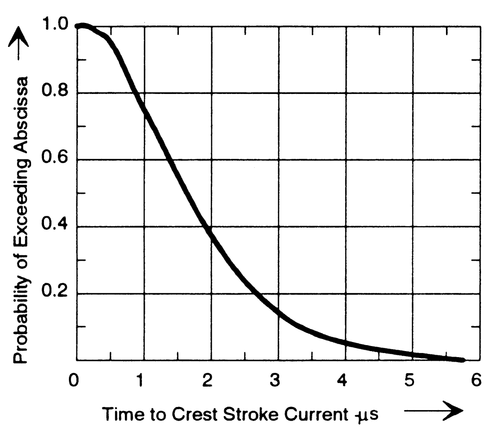


FIGURE 27-12 Distribution of lightning current rise times [recorded by Berger (1967) at Mount San Salvatore].

The frequency of occurrence is also a very important parameter. In order to quantify lightning activity, the crude measure of *thunderstorm day* has been introduced. A thunderstorm day is defined as a 24-h period in which at least one thunder clap has been heard. Collection of historical thunderstorm activity data by the National Weather Service resulted in maps of equi-thunderstorm-day contours. These maps are known as isokeraunic maps, from the Greek word *keraunos*, lightning. Such a map is illustrated in Fig. 27-13. It is important to note that this map, by definition, provides only a crude measure of lightning activity. Specifically, by definition, a thunderstorm day does not provide any information on the frequency of and total lightning activity. Yet, because of lack of better data, the isokeraunic maps have been used for estimation of lightning activity in an area. There are several models, which provide the approximate number

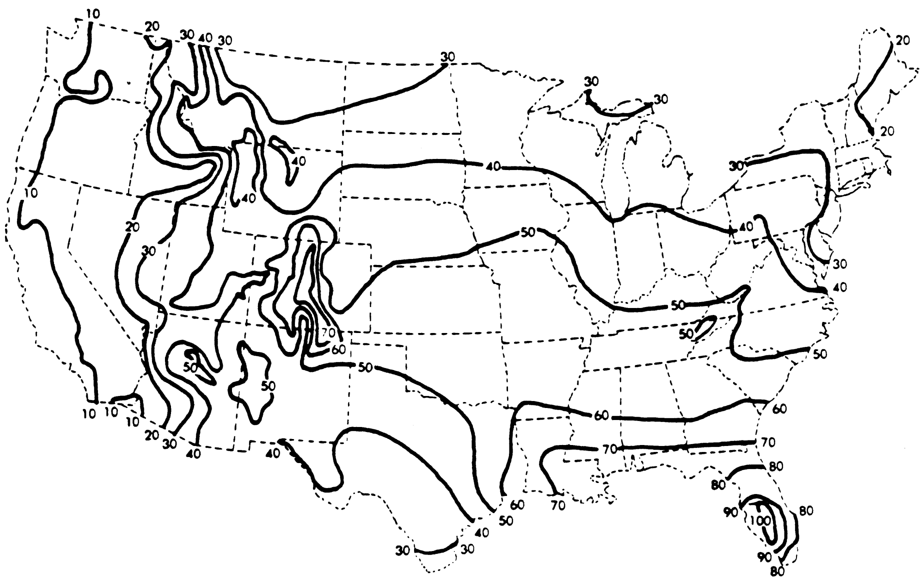


FIGURE 27-13 Isokeraunic curves for the continental United States.

of cloud-to-ground lightning per unit of area as a function of the isokeraunic level. These models will be discussed in detail in Sec. 27.7. As an example, Anderson (1975) has suggested the following

$$N_1 = 0.12T$$

where N_1 is the ground flash density per square kilometer per year and T is the number of thunderstorm days.

Early in the 1980s, the Electric Power Research Institute sponsored a project at the State University of New York at Albany (SUNYA) which resulted in the National Lightning Detection Network (NLDN) records. The system integrated two networks and basically records cloud-to-ground lightning discharges. The objective of the project was to collect lightning data over a period of 10 years, which could be used for lightning protection of power systems. Figure 27-14 illustrates average lightning flashes per square kilometer for the state of Florida. The data were collected over a 5-year period (1985 to 1989). Note that these data correlate reasonably well with the isokeraunic maps data of Fig. 27-13. Similar systems have been installed in many countries.

Summary. This section has described the mechanism of thundercloud formation and electrification and the initiation, mechanism, and characteristics of lightning. Finally, statistical data on lightning parameters were presented. These data are useful for design work.

27.4 POWER SYSTEM OVERVOLTAGES

The causes of power system overvoltages are numerous and the waveforms are complex. It is customary to classify the transients on the basis of frequency content of the waveforms. In this sense, the following three broad categories are defined:

Power frequency overvoltages

Switching overvoltages

Lightning overvoltages

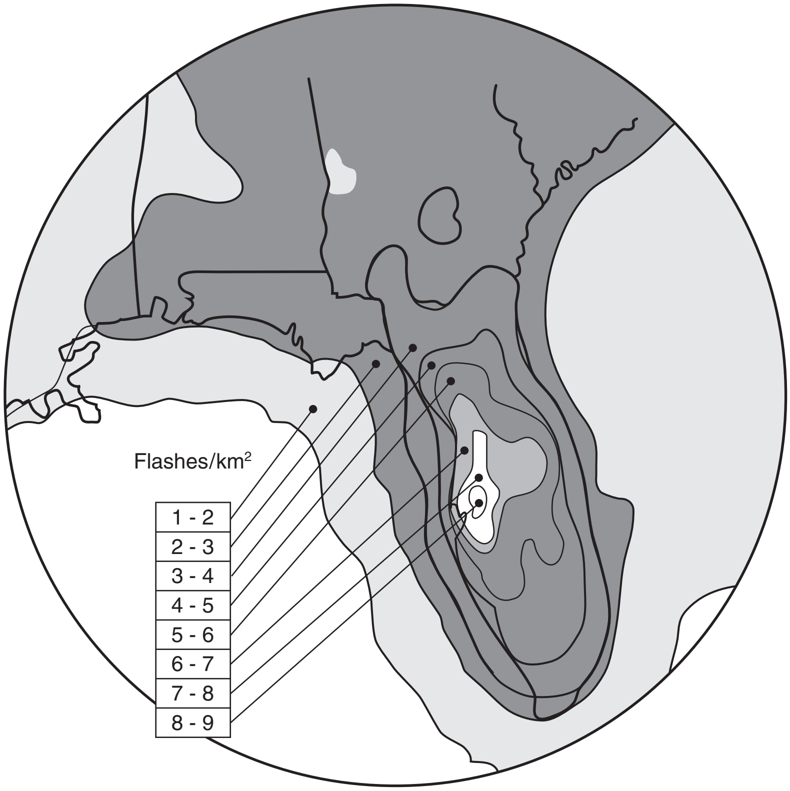


FIGURE 27-14 Average annual flash density data from 1985 to 1989. (*Electric Power Research Institute*.)

Table 27-2 provides brief descriptions and typical causes of the most commonly encountered overvoltages in power systems. The relative level of overvoltages due to these causes is illustrated in Fig. 27-15.

In designing a well-protected electric power system, it is extremely important to thoroughly understand the types, frequency, and magnitude of the expected overvoltages on the power system. For this reason, this section provides a concise discussion of the nature, generation mechanisms, and characteristics of power frequency, switching, and lightning overvoltages in power systems.

Power Frequency Overvoltages. The magnitude of power frequency overvoltages is typically low compared to switching or lightning overvoltages. Specifically, for most causes of these types of overvoltage, the magnitude may be few percent to 50% above the nominal operating voltage. However, they play an important role in the application of overvoltage protection devices. The reason is that modern overvoltage protection devices are not capable of discharging high levels of energy associated with power frequency overvoltages. Thus, it is imperative that protective device ratings be selected in such a way that they do not operate under any foreseeable power frequency overvoltages.

The most common causes of power frequency overvoltages are (1) electric faults, (2) sudden changes of load, and (3) ferroresonance. An electric fault results in voltage collapse for the faulted phase and in a possible overvoltage at the unfaulted phases. The magnitude of the overvoltage depends on the parameters of the circuit, such as positive, negative, and zero sequence impedance, as well as the grounding parameters of the system, such as ground impedance or single- or multiple-grounded

TABLE 27-2 Power-System Overvoltages

Category	Description	Causes
Power frequency overvoltages	Temporary overvoltages dominated by the power frequency component	Electric faults Sudden changes of load Ferroresonance
Switching overvoltages	Temporary overvoltages resulting from a switching operation	Energization of lines Deenergization of capacitor banks Fault interruption/TRV High-speed reclosing Energization/deenergization of transformers Other
Lightning overvoltages	Temporary overvoltages resulting from a lightning stroke terminating at a phase conductor, shield conductor, any other part of a power system, or a nearby object (tree, etc.)	Lightning—cloud-to-ground flashes

system. Figure 27-16 illustrates a typical case of a single-phase-to-ground fault at the end of a 40-mi-long 115-kV transmission line. Because the electric power system is not completely symmetric, the magnitude of the overvoltage on the unfaulted phases may be different; that is, for the case of Fig. 27-16, the overvoltage on phase B is 28.3%, while for phase C the overvoltage is 31.9%. Many studies have

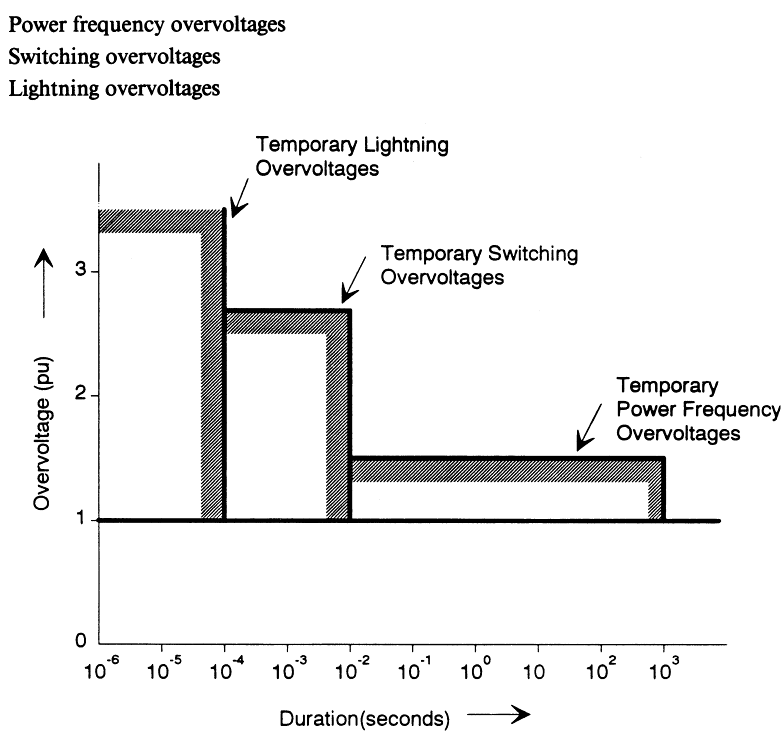


FIGURE 27-15 Typical range of magnitude and duration of power system temporary overvoltages. [*From Regaller (1980).*]

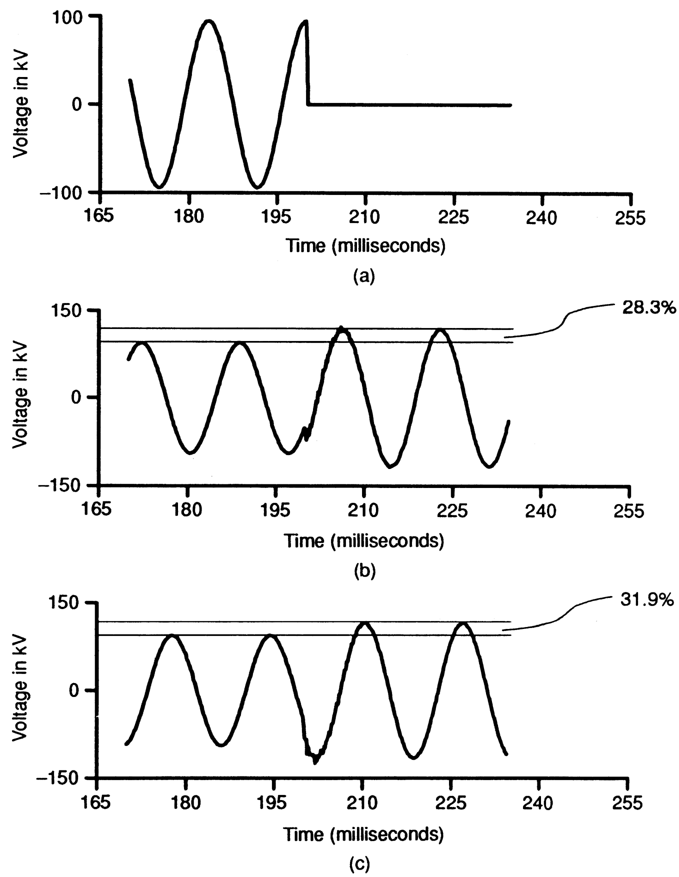


FIGURE 27-16 Overvoltage due to a single-phase-to-ground fault at the end of a 40-mi-long 115-kV line: (*a*) phase A voltage; (*b*) phase B voltage; (*c*) phase C voltage.

been performed over the years to determine simple techniques for determining the power frequency overvoltages. As a first approximation, one can determine the power frequency overvoltage due to a fault from the sequence parameters (positive-, negative-, and zero-sequence impedances) at the fault location. Figure 27-17, taken from Johnson (1979), illustrates the power frequency overvoltage at the unfaulted phases due to a ground fault in one phase as a function of the ratios (X_0/X_1) and (R_0/X_1).

Computer models for determining the power frequency overvoltage by taking into consideration all relevant factors have been developed. Using these models, one can determine the exact power frequency overvoltage and the effect of grounding practices. As an example, Fig. 27-18, taken from Mancao et al. (1992), illustrates the maximum line-to-ground overvoltage, per unit (pu), on typical distribution circuits versus fault distance from the feeding substation.

Another source of power frequency overvoltages is the so-called Ferranti effect, which occurs when a load is disconnected at the end of a long transmission line. In this case, the line draws a capacitive current from the source, which generates a voltage gradient along the line of such a phase as to increase the voltage at the open end of the line. An approximate expression of the overvoltage at the open end of the line is given by

$$\text{Overvoltage in pu} = \frac{1.0}{\cos(\beta l)}$$

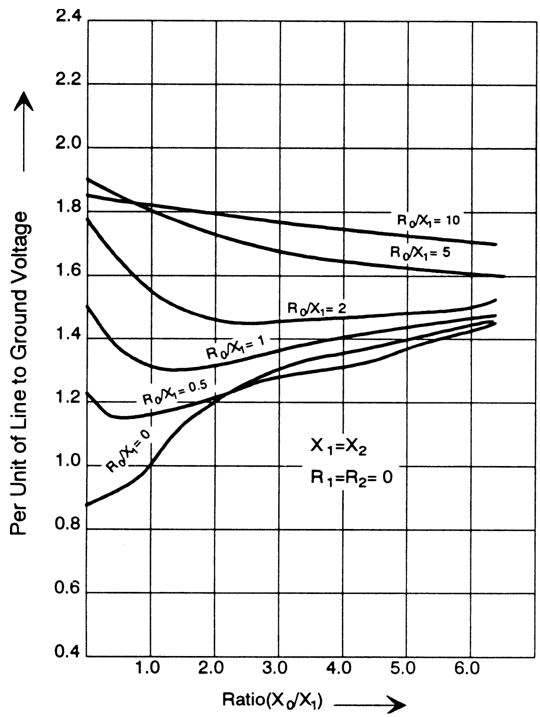


FIGURE 27-17 Overvoltage on unfaulted phase during single-line-to-ground fault. [*From Johnson (1979).*]

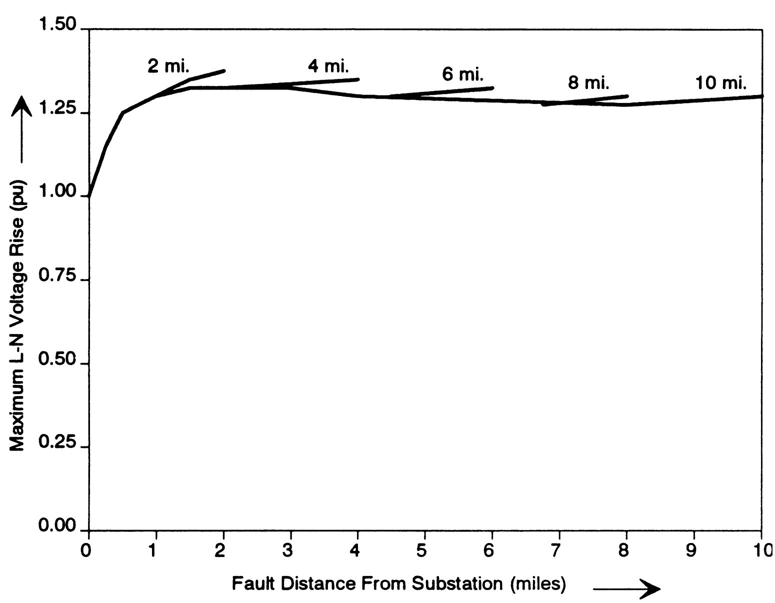


FIGURE 27-18 Overvoltage on unfaulted phases of a distribution circuit as a function of fault distance and circuit length. [*From Mancao et al. (1992).*]

where β is the propagation characteristic of the line ($\omega\sqrt{LC}$) and l is the total length of the line. A typical transmission line of 400 mi may experience an overvoltage of 1.30 pu when one end of the line is open.

Ferroresonance is another cause of power frequency overvoltages but is less frequent. Ferroresonance may occur when energizing long transmission lines and unloaded power transformers, in single-phase switching of a 3-phase transformer, and in other cases involving an iron-core magnetic circuit connected to a substantially capacitive circuit. The overvoltages resulting from ferroresonance can be serious and especially destructive to gapless arresters present in the system. As an example, Fig. 27-19, taken from an IEEE committee report, illustrates the maximum overvoltage due to ferroresonance involving single- or double-phase switching of a 3-phase transformer with an ungrounded primary (delta or wye). The severity of ferroresonance depends on the amount of capacitive reactance present in the system.

Other causes of power frequency overvoltages are (1) generator speedup due to load rejection, (2) generator self-excitation, and (3) malfunction of regulating equipment.

Switching. Switchings in a power system occur frequently. A variety of switchings are performed for routine operations or automatically by control and protection systems. Typical switchings are as follows:

Lines (transmission or distribution)

Cables

Shunt/series capacitors

Shunt reactors

Transformers

Generators/motors

Another class of switching transients are those generated from *insulation flashovers* and breaker *restrikes*. These phenomena are equivalent to the closing of a switch and generate switching surges, which propagate in the system.

Overvoltages resulting from switching operations are typically proportional to the power frequency voltage. For example, energization of a 3-phase line can result in an overvoltage at the open end, which can be as high as 5 pu, depending on the timing of switching with respect to the source. The frequency

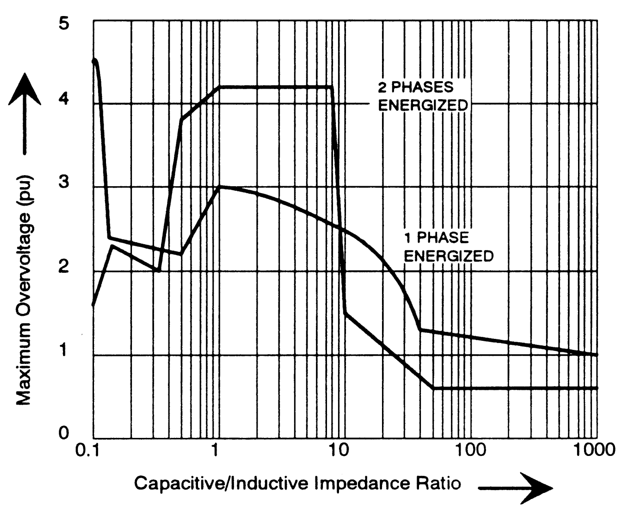


FIGURE 27-19 Maximum overvoltage due to ferroresonance triggered by single-or double-phase switching of a 3-phase transformer with an ungrounded primary.

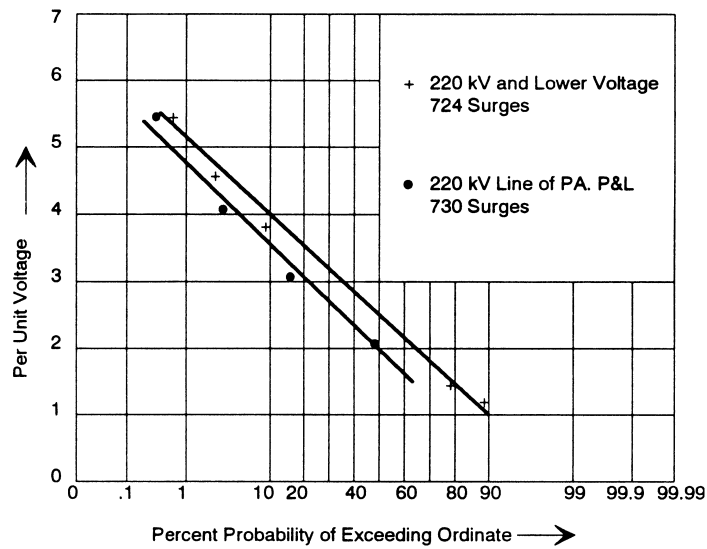


FIGURE 27-20 Probability distribution curves of measured switching overvoltages. [*From Johnson (1979).*]

content of switching transients depends on system parameters. As an example, Fig. 27-20, taken from Johnson (1979), illustrates probability distribution curves of measured line switching overvoltages. Note that there is a substantial probability for overvoltages higher than 5.0 pu.

Switching transients for extra-high-voltage systems, that is, 230 kV and above, can be quite high and must be controlled to avoid the need for higher insulation. There are two methods for controlling the magnitude of switching overvoltages: (1) using breakers with resistor preinsertion and (2) using opening resistors or wound-type potential transformers to discharge trapped charge on lines.

Breakers with resistor preinsertion place a resistor between source and line under energization for a short duration (e.g., 0.8 ms) prior to a direct connection of the source to the line. Proper selection of resistor values and insertion time enables effective control of maximum switching overvoltages. As an example, Fig. 27-21 illustrates the probability distribution curve of switching overvoltages when a resistor preinsertion breaker is used. Note that the maximum switching overvoltage is below 2.5 pu.

Trapped charge on a transmission line can cause excessive switching overvoltages (for specific timing of line energization with respect to source phase). In addition, trapped charge can cause breaker restrike because it contributes to overstressing the breaker insulation. Wound-type potential transformers or opening resistors provide a mechanism for quick drainage of trapped charge on lines.

Switchings can cause other undesirable effects such as inrush currents in transformers and ferroresonance, which has already been discussed.

Lightning Overvoltages. Electric power systems are exposed to weather and therefore are subjected to lightning strikes, which result in overvoltages. Lightning overvoltages are generated by direct lightning strikes on a power system apparatus or indirect strikes to nearby objects, from which subsequent overvoltage is transferred to the system via inductive, capacitive, and conductive coupling.

Unlike power frequency overvoltages and switching overvoltages, which are proportional to the system voltage, lightning overvoltages are independent of system voltage but depend on system impedances. For example, a direct lightning hit to a phase conductor of an overhead transmission line will generate an overvoltage proportional to the characteristic impedance of the line and proportional to the current magnitude of the lightning stroke. This overvoltage may be several million volts. It is a practical and economical impossibility to insulate distribution or lower-kilovolt-level

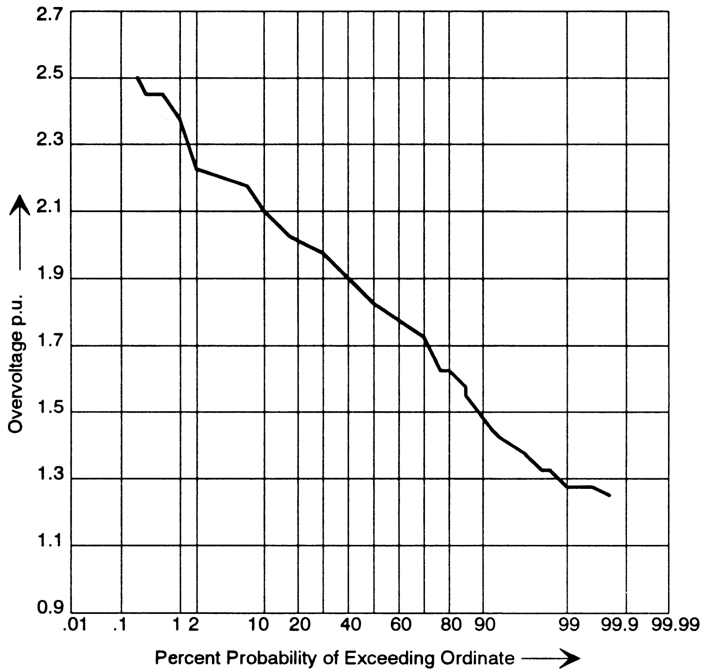


FIGURE 27-21 Probability distribution curves of computed switching overvoltages (switching with a resistor preinsertion breaker).

transmission lines (i.e., 230 kV and below) to withstand this type of overvoltage. As will be discussed in Sec. 27.7, a coordinated design procedure is applied to minimize the effects of lightning; this procedure involves among other things: (1) shielding of lines and equipment, (2) effective grounding, and (3) application of protective devices (surge arresters). The presence of the shielding system ensures that lightning, which otherwise will terminate to a phase conductor, will terminate on a wire, terminal, etc., which is electrically connected to the grounding system. A well-designed grounding system will divert the majority of the lightning stroke current into the soil and thus will minimize the destructive lightning overvoltages. The subject will be further discussed later. Here, a typical example of lightning overvoltages on a 115-kV shielded line is shown in Fig. 27-22. The figure shows the overvoltage at the top of the tower, voltage across insulator of phase A, and the ground potential rise at the tower base. The voltages are given in kilovolts per kiloampere of lightning stroke current. The case shown is for a relatively short tower with effective grounding. The figure shows the effects of nearby towers, which generate reflections of the lightning surge with the end effect of quickly reducing the lightning overvoltages. It should be apparent that the tower-grounding system plays an important role in determining the magnitude of lightning overvoltages, as illustrated in Fig. 27-23. The topic of grounding will be further discussed later.

Lightning strokes to nearby trees, ground, or other objects can result in voltage surges into the power system through coupling. The coupling can be conductive through the conductive soil and the power system grounding structures, inductive, or capacitive. In a typical situation, all the coupling mechanisms may be present, resulting in a voltage surge to the power system. These voltages are called *induced voltage surges* and are generally much lower than those occurring after a direct strike. Specifically, they rarely exceed 400 kV. The induced lightning overvoltages are of concern for distribution lines 35 kV or below. Higher-kilovolt-level lines (i.e., 69 kV and above) have sufficient insulation withstand so that induced lightning voltages do not present the risk of flashover.

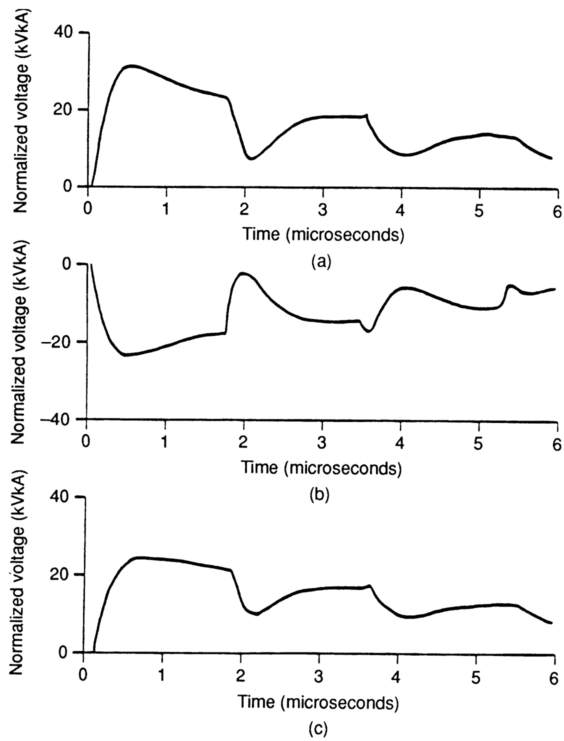


FIGURE 27-22 Typical lightning overvoltages on a transmission line: (*a*) top of tower voltage; (*b*) voltage across insulator (phase A); (*c*) tower ground potential ruse. [*From Meliopoulos (1988).*]

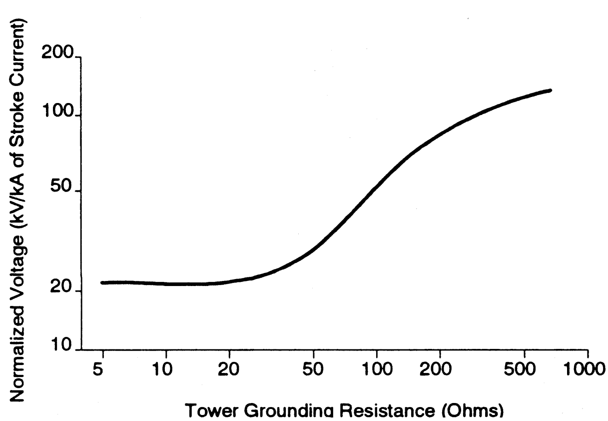


FIGURE 27-23 Effects of tower footing resistance on a specific 115-kV transmission line (standard lightning wave = 1.2/50 ms^{-1}). [*From Meliopoulos (1988).*]

27.5 ANALYSIS METHODS

Design of overvoltage protection systems requires a thorough understanding and analysis of transient overvoltages in power systems. Over the years, many analysis methods have been developed for this purpose. All analysis methods require a proper model of the system under study. In this section, we shall discuss modeling requirements for transient analysis and various analysis methods.

Modeling is probably the most important task in a study. There are many modeling choices which must be made in a prudent way and in view of the objectives of the study. Modeling choices are affected by

Phenomenon under study

Period of concern

Model selection of individual system components

For overvoltage analysis, typical phenomena under study will be line switching, capacitor bank switching, and lightning. The period of concern may be seconds, milliseconds, or microseconds. Model selection of individual system components should be guided by the expected frequency content of the transient. The selected models should have the proper frequency response required for the study under consideration. There is a large number of components to be modeled. A representative list is given in Table 27-3.

TABLE 27-3 A Representative List of Power-System Components

Transmission lines
Single-phase
Three-phase
Overhead
Underground
Lumped capacitors
Iron-core transformers
Generators
Surge arresters
Grounding
Switches, fuses, etc.

Two other related issues are (1) what to model and (2) how to model. The types of choices to be made regarding the question of what to model may include (1) bus inductance, (2) bus capacitance, (3) transformer winding capacitance, and (4) separation distance between arrester and transformer. The question of how to model is complicated. Typical choices are illustrated in Table 27-4. The task of component model selection is very important. To illustrate the point, consider a 40-mi-long transmission line. The line can be represented as a distributed-parameter model or a lumped-parameter model consisting of a set of cascaded pi sections. Figure 27-24 illustrates the switching voltages computed for this line by using the two models. Note that the solutions are different. Of course, the distributed-parameter model provides the correct answer. Examples of simplified lumped-parameter models and distributed-parameter models are illustrated in Fig. 27-25.

TABLE 27-4 Typical Model Choices for Components

Component description	Mathematical model
Lumped-parameter model	Ordinary differential equations
Resistors	
Capacitors	
Inductors	
Distributed-parameter model	Partial differential equations
Lines	
Buses	
Iron-core transformers	Nonlinear equations
Surge arresters	Nonlinear/time-varying equations
Circuit-breaker operator	Logical equations

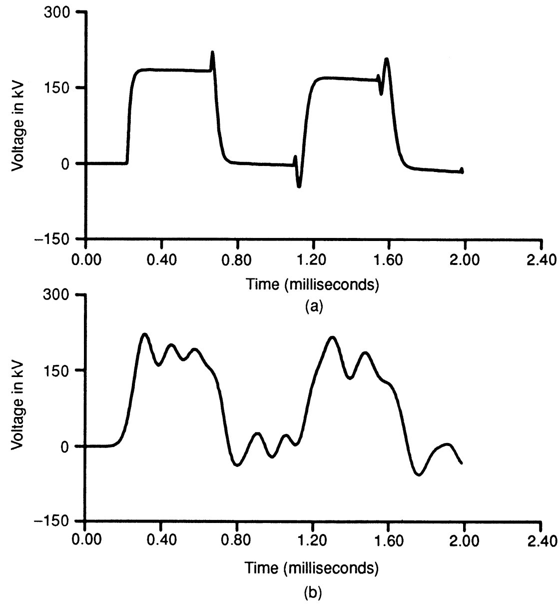


FIGURE 27-24 Switching overvoltage at the open end of a 40-mi-long 115-kV line: (*a*) distributed-parameter model; (*b*) model with six cascaded pi sections.

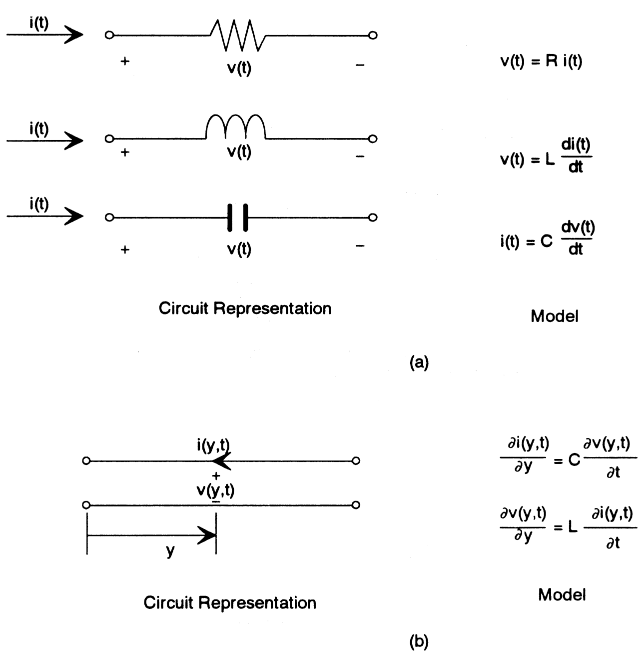


FIGURE 27-25 Power system models for transient analysis: (*a*) lumped-parameter models; (*b*) distributed-parameter models. [*From Meliopoulos (1988).*]

Once the model has been selected, analysis can determine the transient overvoltages for a specific event. Analysis methods can be classified into three categories:

Graphical methods
Analytical methods
Numerical methods

Graphical Methods. The graphical method is based on the observation that the solution for transient voltages and currents in a transmission line can be represented with traveling waves along the line. Under the assumption of an ideal transmission line (zero losses and constant inductance and capacitance per unit length), the waves travel along a line without distortion. Figure 27-26 illustrates the general solution in this sense. Waves are altered when they reach a discontinuity. Specifically, at discontinuity, a wave will be partially reflected and partially transmitted. If the discontinuity involves only resistive elements, the coefficients of reflection and transmission will be constant. This situation is illustrated in Fig. 27-27. The basic relationship among incident (subscript i), reflected (subscript r), and transmitted (subscript t) waves are as follows:

Reflected wave:

$$E_r = \alpha E_i$$

$$I_r = -\alpha I_i$$

$$\alpha = \frac{Z_{\text{eq}} - Z_{01}}{Z_{\text{eq}} + Z_{01}} \quad \text{(reflection coefficient)}$$

$$Z_{\text{eq}} = \frac{R_L Z_{02}}{R_L + Z_{02}}$$

Transmitted wave:

$$E_t = \delta E_i$$

$$I_t = \delta \frac{Z_{01}}{Z_{02}} I_i$$

$$\delta = \frac{2Z_{\text{eq}}}{Z_{\text{eq}} + Z_{01}} \quad \text{(transmission coefficient)}$$

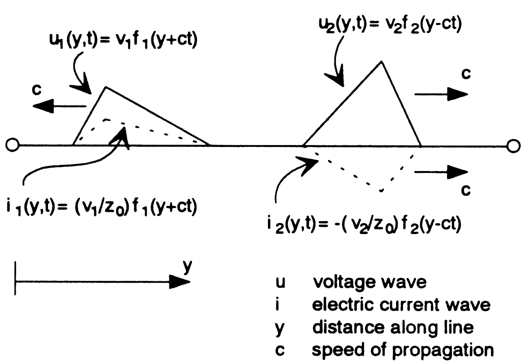


FIGURE 27-26 General wave solution for an ideal distributed-parameter single-phase transmission line.

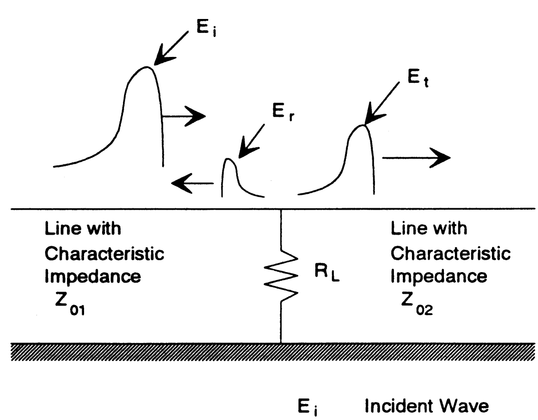


FIGURE 27-27 Transmission and reflection of waves at discontinuity.

If the discontinuity involves storage elements, that is, capacitors or inductors, the coefficients are not constant and the analysis becomes much more complex. The graphical method consists of monitoring all traveling waves on a line with the aid of a diagram, known as the *Bewley diagram*. The Bewley diagram provides, for every point in a system, all the waves present and the time at which they arrive. From this information, the actual voltage waveform at a specific point can be constructed as the superposition of all waves at that point. Such a construction is illustrated in Fig. 27-28.

Analytical Methods. Analytical methods are based on systematic algorithms for solution of the differential equations describing a system. A useful method is based on Laplace transforms. This

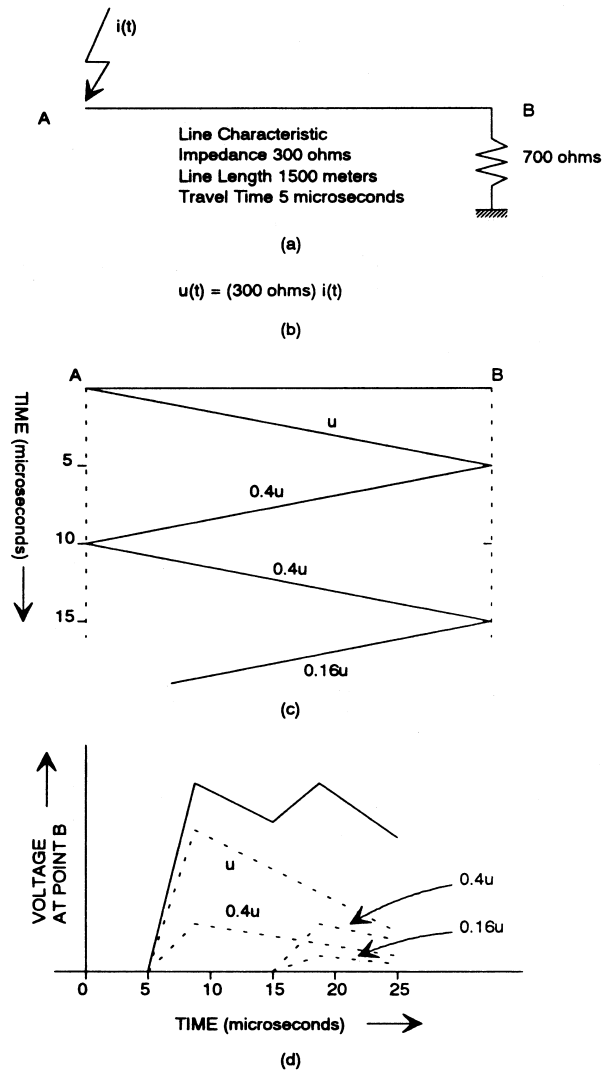


FIGURE 27-28 Illustration of the graphical method: (*a*) system description; (*b*) voltage surge due to lightning *I*(*t*); (*c*) Bewley's diagram; (*d*) construction of voltage at point *B* as the superposition of all surges arriving at point *B*.

method transforms the differential equations describing a component into an equivalent circuit. An example follows. Consider the equation describing an inductor

$$v(t) = L\frac{di(t)}{dt}$$

Application of the Laplace transform on this equation yields

$$V(s) = sLI(s) - Li(0)$$

where $V(s)$ is the Laplace transform of $v(t)$ and $I(s)$ is the Laplace transform of $i(t)$.

The above equation represents the equivalent circuit of Fig. 27-29 (b3). This is known as the Thévenin form of the Laplace domain equivalent circuit. The process can be applied to any power system element described with a set of differential equations. Figure 27-29 illustrates the equivalent circuits in the Laplace domain of typical elements. Note that the equivalent circuits are represented with algebraic equations in complex variables.

Application of the analytical method involves transformation of the differential equations describing individual element with the Laplace transform into an equivalent circuit. Subsequently, nodal analysis (or loop analysis) is applied on the transformed elements to obtain the solution of the voltage at a point of interest as a function of the Laplace variable. Finally, application of the inverse Laplace transform will provide the time waveform of the voltage of interest. Many efficient algorithms were developed during 1975–1995 or so. The details can be found in the literature.

Numerical Methods. Numerical methods are based on transforming the differential equations describing a component into a discrete time equation. This transformation is achieved by proper integration of the differential equations. Many different integration methods can be applied. A very successful method is based on the trapezoidal integration method, because this method is an absolutely stable numerical method. The basic idea is explained as follows. Consider the differential equation

$$\frac{dx(t)}{dt} = ax(t)$$

Integration of this equation in the time interval $(t - h, t)$ yields

$$x(t) - x(t-h) = a\int_{t-h}^{t} x(\tau)d\tau$$

The integral on the right-hand side can be evaluated, *assuming that the function x(t) varies linearly in the time interval (t − h, t)*, yielding

$$\int_{t-h}^{t} x(\tau)d\tau = \frac{h}{2}[x(t) + x(t-h)]$$

This integration is graphically illustrated in Fig. 27-30, which shows that the value of the integral is approximated with the area of the shown trapezoid—thus the name *trapezoidal integration*. Combining above equations and solving for $x(t)$

$$x(t) = \frac{(1 + ah/2)}{(1 - ah)/2} \cdot x(t-h)$$

If $x(0)$ is known, then the above equation can be applied to obtain the value $x(h)$, then $x(2h)$, $x(3h)$, etc. This is a simple algorithm useful for computing the solution at specified times, h, $2h$, $3h$, ...

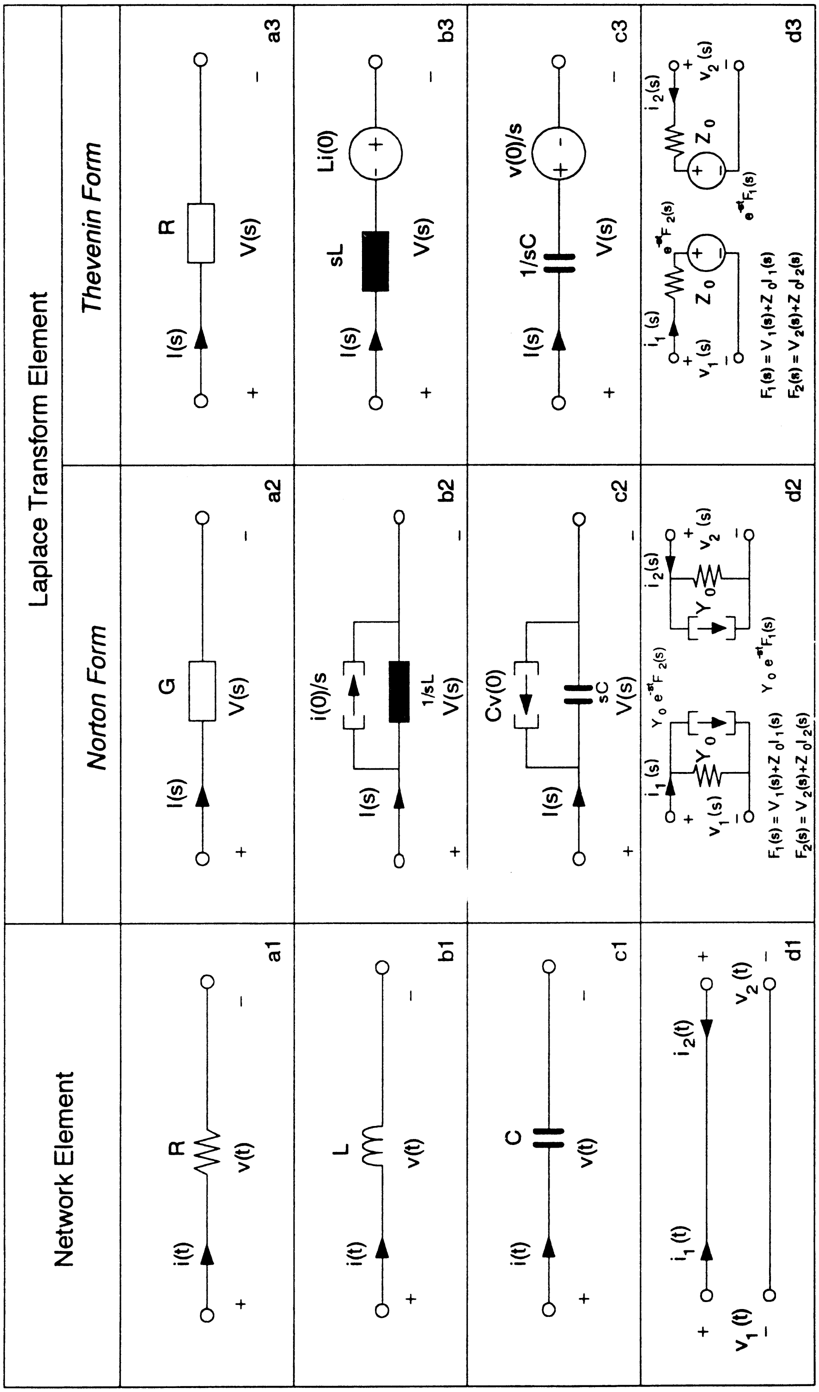


FIGURE 27-29 Equivalent-circuit representation in the Laplace domain. [*From Meliopoulos (1988).*]

The basic idea described above can be applied to the differential equations of any component. The result will be a set of algebraic equations which can be interpreted as a *resistive companion circuit*. The results for simple elements are illustrated in Fig. 27-31. Application of the method requires transformation of each element of the power system into a resistive companion circuit. The process replaces the actual system with a resistive network. This network includes voltage and current sources that depend on the state of the system at times less than t. Application of nodal analysis (or loop analysis) provides the solution for voltages and currents at time t. The process can start at time $t = 0$ and be repeated at times $t = h, t = 2h, \dots$, yielding the values of voltages and currents at times $t = 0, h, 2h, \dots$, etc. Further details can be found in Meliopoulos (1988).

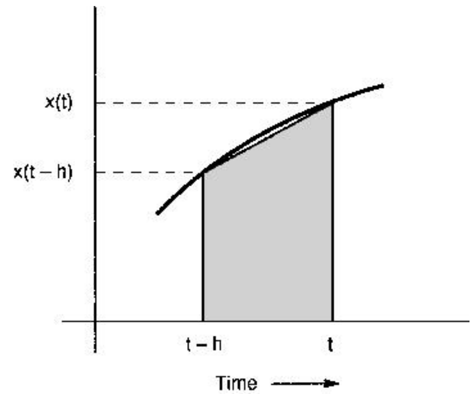


FIGURE 27-30 Graphical representation of the trapezoidal integration.

3-Phase Transmission Lines. 3-Phase transmission is represented with a distributed-parameter model. This model can be derived by considering an infinitesimal length of a 3-phase line as in Fig. 27-32*b*. Assuming an ideal line (zero losses and constant inductance and capacitance per unit length), the model equations are

$$CL\frac{\partial^2}{\partial t^2}v_{abc}(y, t) = \frac{\partial^2}{\partial y^2}v_{abc}(y, t)$$

$$CL\frac{\partial^2}{\partial t^2}i_{abc}(y, t) = \frac{\partial^2}{\partial y^2}i_{abc}(y, t)$$

$$L\frac{\partial}{\partial t}i_{abc}(y, t) = \frac{\partial}{\partial y}v_{abc}(y, t)$$

where

$$v_{abc}(y, t) = \begin{bmatrix} v_a(y, t) \\ v_b(y, t) \\ v_c(y, t) \end{bmatrix}$$

$$i_{abc}(y, t) = \begin{bmatrix} i_a(y, t) \\ i_b(y, t) \\ i_c(y, t) \end{bmatrix}$$

$$L = \begin{bmatrix} L_{aa} & L_{ab} & L_{ac} \\ L_{ab} & L_{bb} & L_{bc} \\ L_{ac} & L_{bc} & L_{cc} \end{bmatrix}$$

is the inductance matrix per unit length of the line and

$$C = \begin{bmatrix} C_{aa} & C_{ab} & C_{ac} \\ C_{ab} & C_{bb} & C_{bc} \\ C_{ac} & C_{bc} & C_{cc} \end{bmatrix}$$

is the capacitance matrix per unit length of the line, y is distance along line, and t is time.

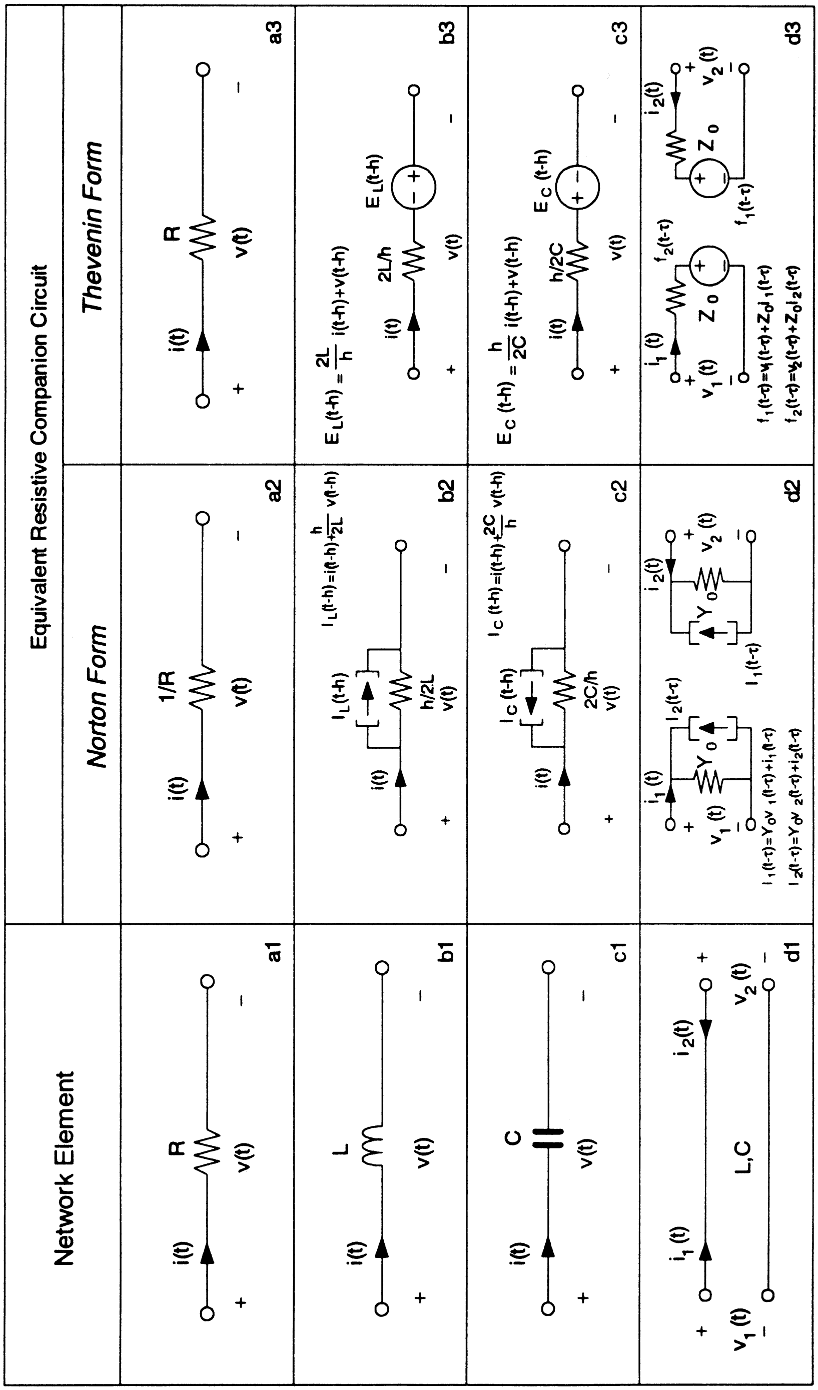


FIGURE 27-31 Equivalent resistive comparison circuits. [*From Meliopoulos (1988).*]

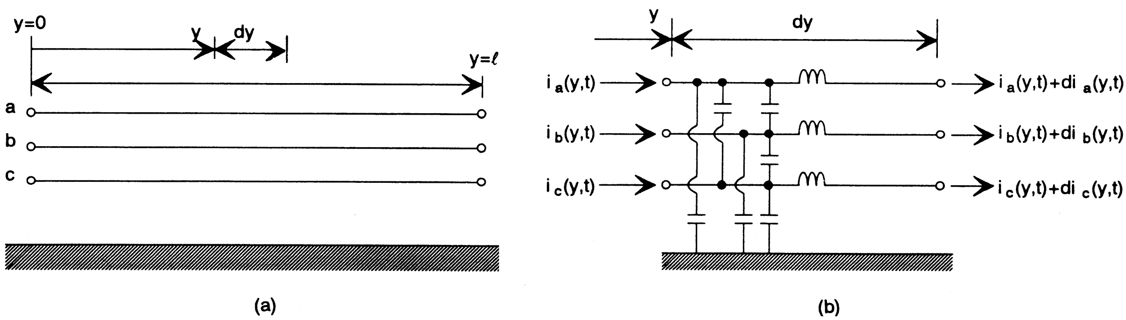


FIGURE 27-32 3-Phase line model: (*a*) 3-phase transmission line; (*b*) equivalent circuit of an infinitesimal length (*dy*) of 3-phase line.

This is a rather simplified model of a 3-phase line and yet very complex. To make clear the behavior of the 3-phase line, the concept of the *ideal continuously transposed line* will be introduced. This concept assumes a perfectly symmetrical line

$$L_{aa} = L_{bb} = L_{cc} = L_s$$

$$L_{ab} = L_{ac} = L_{bc} = L_m$$

$$C_{aa} = C_{bb} = C_{cc} = C_s$$

$$C_{ab} = C_{ac} = C_{bc} = C_m$$

Next, Karrenbauer's transformation is introduced as follows

$$v_{abc}(y, t) = Kv_{gll}(y, t)$$

$$i_{abc}(y, t) = Ki_{gll}(y, t)$$

where

$$v_{gll}(y, t) = \begin{bmatrix} v_g(y, t) \\ v_{l1}(y, t) \\ v_{l2}(y, t) \end{bmatrix}$$

$$i_{gll}(y, t) = \begin{bmatrix} i_g(y, t) \\ i_{l1}(y, t) \\ i_2(y, t) \end{bmatrix}$$

and

$$K = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

Replacement of the actual voltages and currents $v_{abc}(y, t)$ and $i_{abc}(y, t)$ with the voltages and currents $v_{gll}(y, t)$ and $i_{gll}(y, t)$ through Karrenbauer's transformation yields the following transformed equations for the 3-phase line (Meliopoulos 1988)

Set 1 (ground-mode equations)

$$(L_s + 2L_m)(C_s + 2C_m)\frac{\partial^2}{\partial t^2}v_g(y, t) = \frac{\partial^2}{\partial y^2}v_g(y, t)$$

$$(L_s + 2L_m)(C_s + 2C_m)\frac{\partial^2}{\partial t^2}i_g(y, t) = \frac{\partial^2}{\partial y^2}i_g(y, t)$$

$$(L_s + 2L_m)\frac{\partial}{\partial t}i_g(y, t) = \frac{\partial}{\partial y}v_g(y, t)$$

Set 2 (line-mode 1 equations)

$$(L_s - L_m)(C_s - C_m)\frac{\partial^2}{\partial t^2}v_{l1}(y, t) = \frac{\partial^2}{\partial y^2}v_{l1}(y, t)$$

$$(L_s - L_m)(C_s - C_m)\frac{\partial^2}{\partial t^2}i_{l1}(y, t) = \frac{\partial^2}{\partial y^2}i_{l1}(y, t)$$

$$(L_s - L_m)\frac{\partial}{\partial t}i_{l1}(y, t) = \frac{\partial}{\partial y}v_{l1}(y, t)$$

Set 3 (line-mode 2 equations)

$$(L_s - L_m)(C_s - C_m)\frac{\partial^2}{\partial t^2}v_{l2}(y, t) = \frac{\partial^2}{\partial y^2}v_{l2}(y, t)$$

$$(L_s - L_m)(C_s - C_m)\frac{\partial^2}{\partial t^2}i_{l2}(y, t) = \frac{\partial^2}{\partial y^2}i_{l2}(y, t)$$

$$(L_s - L_m)\frac{\partial}{\partial t}i_{l2}(y, t) = \frac{\partial}{\partial y}v_{l2}(y, t)$$

Note that the complex equations for the 3-phase line have been replaced with three sets of equations, each set representing an ideal single-phase line. The characteristic impedance and speed of propagation of surges for the three ideal single phase lines are

Ground mode

$$Z_g = \frac{1}{Y_g} = \sqrt{\frac{L_s + 2L_m}{C_s + 2C_m}}$$

$$c_g = \frac{1}{\sqrt{(L_s + 2L_m)(C_s + 2C_m)}}$$

Line mode (1 or 2)

$$Z_l = \frac{1}{Y_l} = \sqrt{\frac{(L_s - L_m)}{(C_s - C_m)}}$$

$$c_l = \frac{1}{\sqrt{(L_s - L_m)(C_s - C_m)}}$$

This model of the line is illustrated in Fig. 27-33. The three sets of equations or three ideal single-phase line models above are known as (1) the ground mode g, (2) line mode 1 l_1, and (3) the line mode 2 l_2. The names become obvious if one considers excitation of the 3-phase line with one mode.

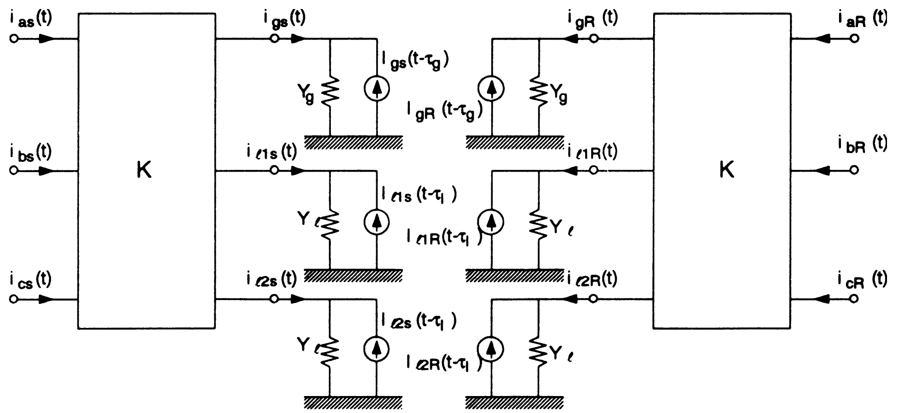


FIGURE 27-33 Equivalent circuit of a continuously transposed line.

For example, consider that the line is excited in such a way that $i_{l1}(y, t) = 0$ and $i_g(y, t) = 0$, $i_{l2}(y, t) = 0$. In this case, the actual phase currents will be

$$i_{abc}(y, t) = Ki_{gll}(y, t)$$

which yields

$$i_a(y, t) = i_{l1}(y, t)$$
$$i_b(y, t) = -2i_{l1}(y, t)$$
$$i_c(y, t) = i_{l1}(y, t)$$

This state of the line is illustrated in Fig. 27-34*a*. Note that there is a positive surge on phases *a* and *c*, which returns through phase *b*. All electric current surges are confined in the line conductors, thus the name line mode.

Consider now excitation with $i_g(y, t) \neq 0$, $i_{l1}(y, t) = 0$, and $i_{l2}(y, t) = 0$. In this case, the equation

$$i_{abc}(y, t) = Ki_{gll}(y, t)$$

yields

$$i_a(y, t) = i_g(y, t)$$
$$i_b(y, t) = i_g(y, t)$$
$$i_c(y, t) = i_g(y, t)$$

This state of the line is illustrated in Fig. 27-34*b*. Note that the surges on all phases are equal. The return current is through the earth, thus the name ground mode. The parameters of the various modes of propagation of surges in a 3-phase line are different. As an example, the following values apply to a 115-kV 3-phase line.

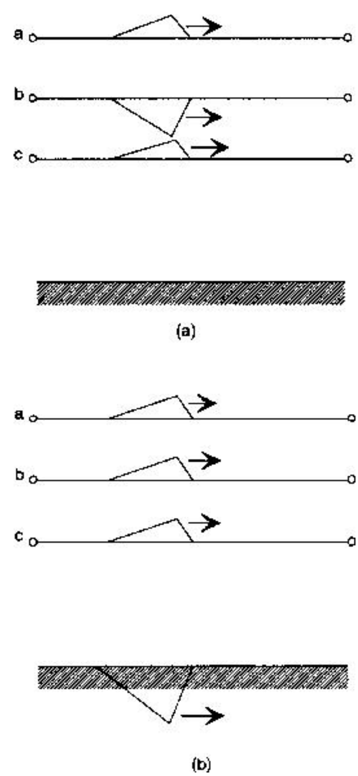


FIGURE 27-34 Modes of propagation along a 3-phase line: (*a*) line mode; (*b*) ground mode.

The inductance matrix is

$$L = \begin{bmatrix} 2.2289 & 1.032 & 0.8935 \\ 1.032 & 2.2289 & 1.032 \\ 0.8935 & 1.032 & 2.2289 \end{bmatrix} \quad \mu\text{H/m}$$

Therefore $\qquad L_s = 2.2289 \quad \mu\text{H/m} \qquad L_m = 0.9858 \quad \mu\text{H/m}$

The capacitance matrix is

$$C = \begin{bmatrix} 7.6737 & -1.9129 & -1.0065 \\ -1.9129 & 8.0184 & -1.9129 \\ -1.0065 & -1.9129 & 7.6737 \end{bmatrix} \times 10^{-12} \quad \text{F/m}$$

Therefore $\qquad C_s = 7.7886 \times 10^{-12}\ \text{F/m} \qquad C_m = -1.608 \times 10^{-12}\ \text{F/m}$

Line mode (1 or 2):

$$Z_l = \sqrt{\frac{(L_s - L_m)}{(C_s - C_m)}} = 363.7\ \Omega$$

$$c_l = \frac{1}{\sqrt{(L_s - L_m)(C_s - C_m)}} = 2.926 \times 10^8\ \text{m/s}$$

Ground mode:

$$Z_g = \sqrt{\frac{(L_s + 2L_m)}{(C_s + 2C_m)}} = 959.1\ \Omega$$

$$c_g = \frac{1}{\sqrt{(L_s + 2L_m)(C_s + 2C_m)}} = 2.283 \times 10^8\ \text{m/s}$$

When lightning hits a 3-phase line, all modes of propagation are excited. The overvoltage of the 3-phase line as a result of the lightning stroke is determined from the parameters of all the modes. Later, examples of these calculations will be provided.

Frequency-Dependent Models. The parameters of power-system elements are frequency dependent. As an example, consider the ground-mode resistance of a 3-phase line. This resistance value may change *several orders of magnitude* in the frequency range 60 Hz (power frequency) to 1 MHz. This increase of resistance is mainly due to skin-effect-type phenomena. The higher resistance values at high frequencies tend to attenuate higher-frequency components of transient overvoltages with the end effect of reducing the maximum overvoltages. Computer models to account for the frequency dependence have been developed for almost all power components. The reader is referred to the literature. Frequency-dependent models are complex, but, on the other hand, they provide a better representation of the real system. As an example, consider Fig. 27-35. The figure illustrates comparison of measured and computed switching transients in a 224.15-mi-long 230-kV line. The single-line diagram of the system is illustrated in Fig. 27-36. The tests were performed by Bonneville Power Administration. The calculated switching transient was performed with a frequency-dependent model (Cokkinides and Meliopoulos 1988). Note that the maximum switching overvoltage is 1.81 pu. It is important to note that simulation of the same system assuming an ideal continuously transposed line results in a maximum switching overvoltage of 2.05 pu. The conclusion is that frequency-dependent models are more realistic.

Grounding Models. Grounding plays an important role in dissipation of lightning strokes and therefore controlling overvoltages resulting from lightning. Yet, grounding has been widely misunderstood and proper analysis models are scarce. Modeling of grounding systems is a rather complex task, and it is strongly coupled to the overall modeling procedure for power systems. Two distinct approaches apply: (1) grounding models for low frequency generally consider grounds as a pure resistance and thus dc analysis models are utilized, based on the method of moments or relaxation methods; (2) grounding models for higher frequencies require complete electromagnetic analysis. For this purpose, finite element analysis or the method of moments can be utilized. Since the grounds are typically complex systems, simplifications are typically introduced. A rule of thumb for selecting dc models or the more complex models is to compare the largest dimension of a grounding system l to the depth of penetration $\delta = \sqrt{2}/\sqrt{\mu\omega\sigma}$. If $l > 0.1\delta$, then a complete model is necessary. In this expression, μ is the permeability of the medium, ω is the frequency in radians per second, and σ is the conductivity.

The choice of the grounding model can be crucial. As an example, consider a grounding system consisting of a counterpoise buried in soil of 100 $\Omega \cdot$ m. The dc ground resistance of this system is 2.29 Ω. The frequency-dependent impedance is much higher for high frequencies and approaches 2.29 Ω for frequencies below 50 kHz. As a result, a lightning discharge through this ground will require a frequency-dependent ground model to accurately predict the transient overvoltages. As an example, Fig. 27-37 shows the differences between the correct model and the dc model for the standard lightning waveform. The example emphasizes the importance of selecting the proper model for the grounding system.

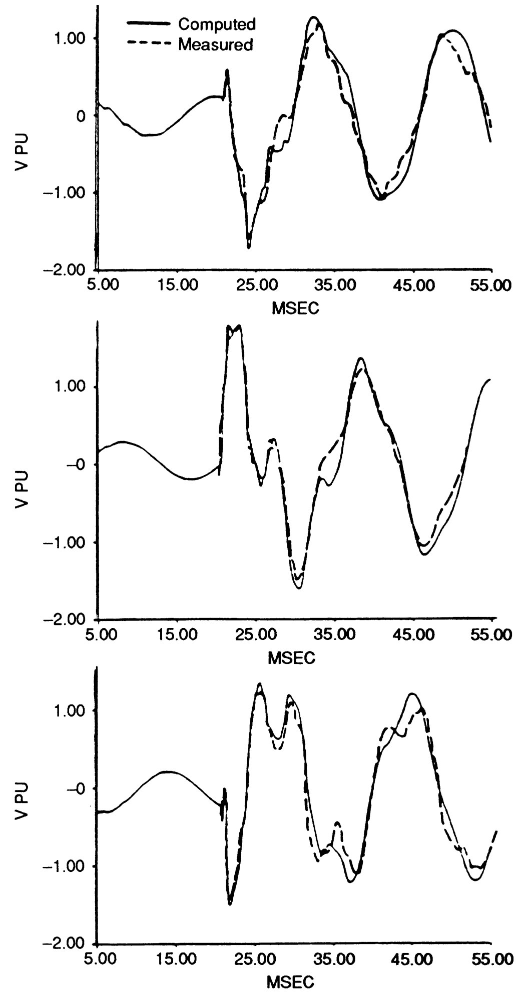


FIGURE 27-35 Comparison of simulation and test results of switching surges: (*a*) phase A transient voltage; (*b*) phase B transient voltage; (*c*) phase C transient voltage. [*From Cokkinides and Meliopoulos (1988).*]

Grounding systems play an important role in propagation of surges from high-voltage power systems to lower-voltage secondary systems. Specifically, transients initiated in the power system can travel through the grounding system and enter the secondary power system of a facility, reaching sensitive electronic equipment. Transformers do not exist in the path of a neutral. Thus, high-frequency transients travel almost unattenuated through the neutral. Attenuation is provided only by the grounds if the neutral is multiply grounded. A proper model of the grounds, neutrals, and the power system can provide a good tool to determine the transient overvoltages reaching a specific piece of equipment. As an example, consider a facility served by a 1-mi (1.6 km) underground distribution cable as in Fig. 27-38. The underground cable is fed from an overhead distribution circuit, which is subject to lightning. At such a service entrance, transients can enter a facility through the cable concentric neutral and the ground conductors.

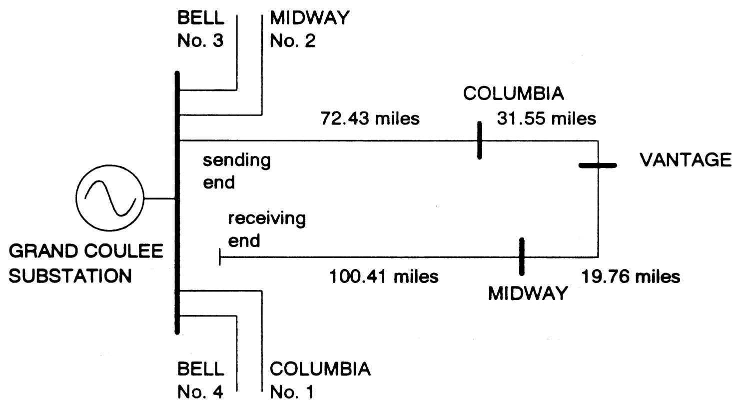


FIGURE 27-36 Illustration of system for switching surge test. *(Bonneville Power Administration.)*

Probabilistic Methods. The objective of these methods is to provide a probabilistic description of overvoltages in a specific power apparatus. The method takes into consideration the uncertainty in parameters affecting the overvoltages. Some of the uncertain parameters are listed in Table 27-5.

The probabilistic methods consist of the following three components:

1. A probabilistic model for the uncertain parameters
2. An analysis model for the system under study
3. A Monte Carlo simulation method

The result of the method is the probability distribution of the overvoltages at the apparatus of interest. Given the probabilistic model for the uncertain parameters and an analysis model for the system under study, the Monte Carlo simulation consists of the following steps:

Step 0: Set count of trial *n* equal to 0.

Step 1: Generate (randomly) a sample of parameters from the probabilistic model of uncertain parameters (crest value of lightning stroke, rise time, fall time, location of incidence, etc.).

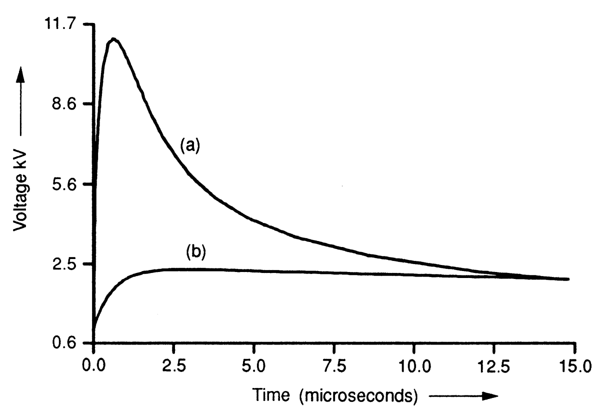


FIGURE 27-37 Transient voltage on a 300-ft counterpoise from a 1.2-50-m/s lightning wave of a 1-kA crest: (*a*) frequency-dependent model; (*b*) frequency-independent model.

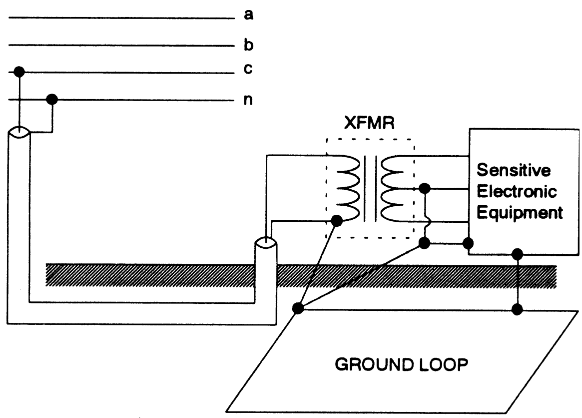


FIGURE 27-38 Illustration of a typical service entrance to a commercial or industrial facility.

Step 2: Use the analysis model and the parameter values determined in step 1 to evaluate the maximum overvoltage at points of interest (or other quantities of interest). Store the computed values.

Step 3: Repeat steps 1, 2, and 3 until the number of trials = n. Then go to step 4.

Step 4: Use the stored computed values of overvoltages (or other quantities of interest) to generate histograms and/or probability distribution functions.

TABLE 27-5 Typical Parameters with Uncertainty

Lightning	Crest value
	Rise time
	Frequency of occurrence
	Shielding failure
	Other
Switching	Switch closing time
	Time of preinsertion of resistors
	Trapped charge
	Other

The utility of probabilistic methods is quite obvious. They provide the tool to determine not only the expected overvoltages but also the frequency of occurrence. For external insulation protection practices, probabilistic methods provide an indispensable tool for optimal designs.

27.6 OVERVOLTAGE PROTECTION DEVICES

With certainty, temporary overvoltages on power-system apparatus will exceed their withstand capability. In this case, if the apparatus is left unprotected, insulation failure will occur. Therefore, it is necessary to protect power system components against overvoltages. The philosophy and objectives of this protection vary depending on the type of overvoltages, frequency, effects of insulation failure, and cost of repair. These issues will be discussed later. In this section, we will be concerned with available protection devices, their characteristics, and protection levels. An overvoltage protection device should ideally limit the voltages across the insulation of a power apparatus below a specified value. This specified value is called the *protection level*. The ideal protection device has the voltage-current characteristic indicated in Fig. 27-39. Specifically, if the voltage across the protection device is less than the protection level, then the protection device should have an infinitely large impedance. If the voltage across the protection device is higher than the protection level, then the protection device should allow the flow of electric current through it

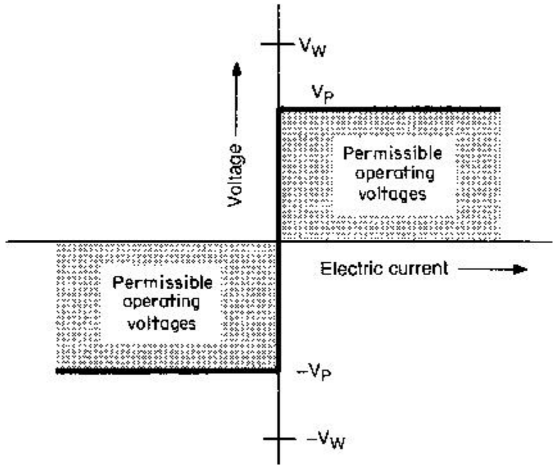


FIGURE 27-39 Voltage-current characteristics of the ideal overvoltage protection device.

in such a way that the voltage is clipped to the protection level value. Such a characteristic can be mathematically described with the equation

$$i(t) = i_0 \left\{ \frac{abs\,[v(t)]}{v_p} \right\}^n \text{sign}\,[v(t)] \tag{27-1}$$

where i_0 = a constant
v_p = protection level
abs $(\cdot)$ = absolute value of the argument $(\cdot)$
$v(t)$ = voltage across the protective device
sign $[v(t)]$ = sign of voltage $v(t)$, + or −
n = a very large number

The protection characteristics of Fig. 27-39 correspond to a very large value of n.

Of course, a protection device with these characteristics does not exist. Research and development has resulted in protection devices which come close to that of Fig. 27-39 to varying degrees. It is important to note that the research and development of protection devices has been evolutionary; over the years the following major breakthroughs have occurred:

Air gaps
Gapped surge arresters
Expulsion arresters
Gapped valve-type arresters
Gapped silicon carbide (SiC)
Metal oxide varistors
Shunt-gapped metal oxide varistors (MOVs)

The most important application characteristics of a protection device are the protection level and the reseal level. These are defined as follows:

Protection level: The maximum voltage, which will be allowed by the device across its terminals.

Reseal level: The maximum voltage below which the protection device will disallow significant electric current through it.

For best results, it is expedient that these two parameters be as close as possible. Observe that for the ideal protection device, the protection level is equal to the reseal level. In order to quantify the protection quality of various technologies of protective devices, the following definition is introduced:

Protection quality index (PQI). The protection-quality index is defined as the ratio of the reseal level over the protection level:

$$\text{PQI} = \frac{V_r}{V_p}$$

Note that the ideal protective device has a protection-quality index of 1. As should be expected, the protective capabilities of protective devices increase with every breakthrough in the technology. Air gaps provide rudimentary protection, gapped SiC surge arresters provide improved protection, and finally MOVs provide even better protection. Within each class of technology, many variations exist. The next paragraphs will discuss the above technologies.

Air (Spark) Gaps. A spark gap was the first protection device to be applied on a power system. It can be constructed with two electrodes placed at a certain distance. The shape of the electrodes varies depending on the application—for example, two wires, a wire and a planar electrode, or two spheres. When the voltage across the air gap exceeds a certain value, an arc will initiate between the electrodes. The voltage across the arc will depend on the current of the arc, the length of the arc, and time. A typical variation of the voltage across the gap is illustrated in Fig. 27-40. The protection level as well as the reseal level, indicated as maximum permissible operating voltage across the gap, are shown. Note that for good protection, the protection quality index (the ratio V_r/V_p) should be maximized. Unfortunately, air gaps exhibit very low protection quality index. In many applications, and in order to improve the protection quality index, the electrodes are shaped in such a way as to force the arc to elongate and therefore increase the voltage reseal value V_r.

Gapped Surge Arresters. Gapped surge arresters consist of a series of spark gaps with or without series blocks of nonlinear resistors, which act as current limiters. The construction of such an arrester is illustrated in Fig. 27-41.

The function of the air gap is to isolate the current limiting block of the surge arrester from the power frequency voltage under normal operating conditions. The air gap is necessary because the current limiting block will fail thermally if subjected to the continuous normal power frequency

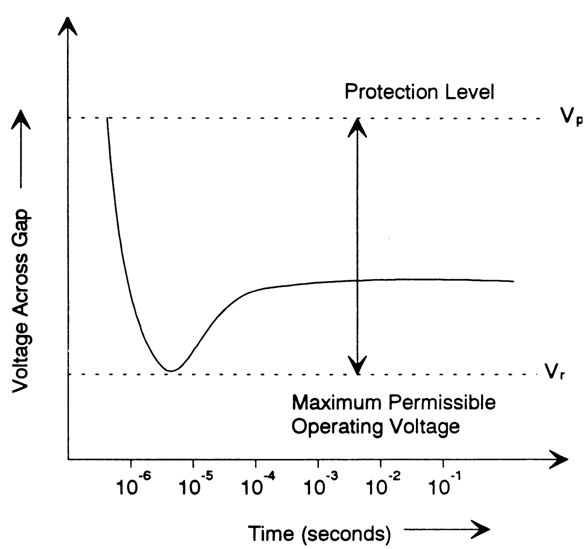


FIGURE 27-40 Typical voltage-time characteristic of a long air gap.

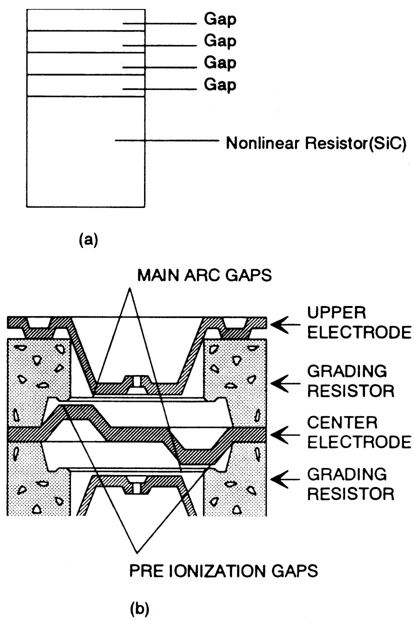


FIGURE 27-41 Construction of a gapped arrester: (*a*) block diagram; (*b*) detail of a gap.

voltage. Elaborate designs of air gaps have been developed, which assure (1) consistent sparkover voltage and (2) arc extinction (resealing) after the overvoltage ceases to exist.

Consistent sparkover voltage is achieved by either a trigger gap or a preionizer. In order to prevent damage of the trigger mechanism, the main current is not allowed to pass through it by use of grading resistors as is illustrated in Fig. 27-41, but through the main arc gaps. To ensure arc extinction and good resealing properties, a mechanism is provided for the control of the arc. Specifically, the main current is passed through a coil. The magnetic field of the coil is established in the space of the arc and is oriented so as to force the arc into a serrated-tooth chamber. The end result of this interaction is that the arc is elongated, cooled, and the voltage across the arc becomes substantial. Figure 27-42 illustrates the buildup of the voltage across the arc due to this mechanism. An illustration of the serrated-tooth chamber is illustrated in Fig. 27-43. The elongated arc improves the capability of the arrester to reseal when the current through it decreases after the overvoltage ceases to exist. It is important to note that the arc control process takes a relatively long time (hundreds of microseconds) to develop full voltage across the gap. This means that for fast transients, that is lightning, almost the entire voltage appears across the current-limiting block.

The current-limiting block in modern gapped arresters is constructed from SiC. The block has nonlinear resistance characteristics that are controlled with the manufacturing process. Specifically, the block is constructed from SiC crystals which are ground into powder form and then are mixed and pressed together with insulating material, forming a block. In this way, an SiC particle partially touches other particles and partially is insulated by the insulating material. The nonlinear properties

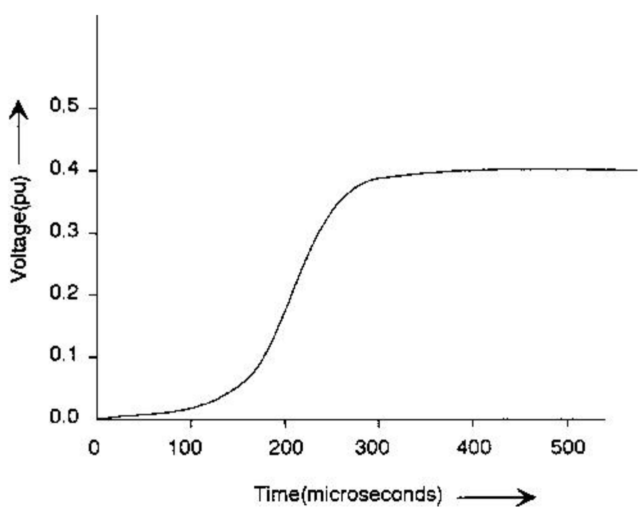


FIGURE 27-42 Evolution of voltage across the arc of a gapped arrester.

are due to the resistance-temperature properties of the junction between the SiC crystals. This relationship is illustrated in Fig. 27-44.

The combined operation of the gap and the current-limiting block result in the voltage-current characteristic illustrated in Fig. 27-45. It is important to note that, while the sparkover voltage V_s is typically consistent, the voltage-current curve depends on the waveform of the electric current, since the gap voltage is time dependent.

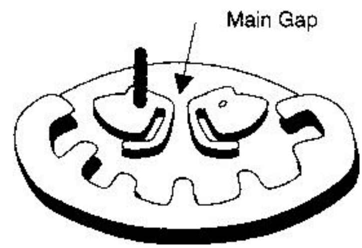


FIGURE 27-43 Illustration of a serrated-tooth chamber of a gapped arrester.

Gapped SiC arresters are subject to power-follow current because of the poor nonlinear characteristics of the SiC block. Specifically, after sparkover and after the temporary overvoltage has subsided, if the power voltage across the arrester is substantial, a relatively high current will continue to flow through the arrester. The reason can be easily visualized if one considers that the arc in the gap is a low-resistance conductor and that the current is primarily determined by the voltage-current characteristics of the SiC block. The situation is illustrated in Fig. 27-46. In the figure, the power-follow current of a SiC gap arrester is also compared to the power-follow current of an MOV arrester, which is discussed next.

Metal Oxide Varistor Arresters. An MOV is formed from a variety of materials via a manufacturing process, which provides the desired electrical properties to the varistor. These electrical properties are not existent in the raw materials. In other words, the electrical properties of the final product are completely dependent on the manufacturing process. Much research effort has been expended to achieve electrical properties of the varistor close to those of the ideal protective device. The typical structure of a metal oxide varistor consists of highly conductive tiny particles of metal oxide (usually zinc oxide, ZnO) suspended in a semiconducting material. This structure is illustrated in Fig. 27-47. The manufacturing process determines the size of the metal oxide particles as well as the thickness and resistivity of the semiconducting material. The key to a good manufacturing process is that the conductive metal oxide particles do not touch each other but are separated by semiconducting material. This structure gives the varistor the properties of a pair of back-to-back-connected zener diodes. A typical voltage-current characteristic of an MOV is illustrated in Fig. 27-46. It should be understood that this function holds for negative voltages and currents.

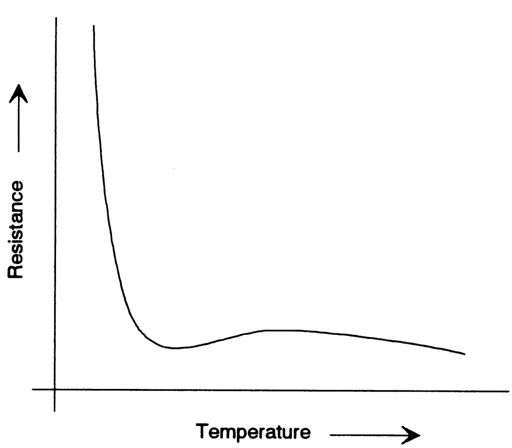


FIGURE 27-44 Typical resistance versus temperature of junction between SiC crystals.

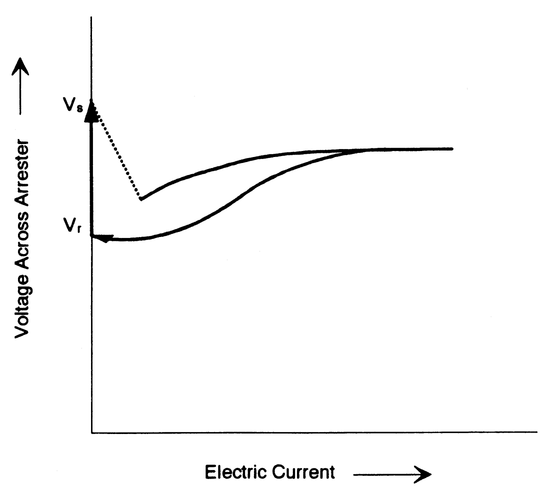


FIGURE 27-45 Typical voltage-current characteristics of a modern SiC gapped arrester.

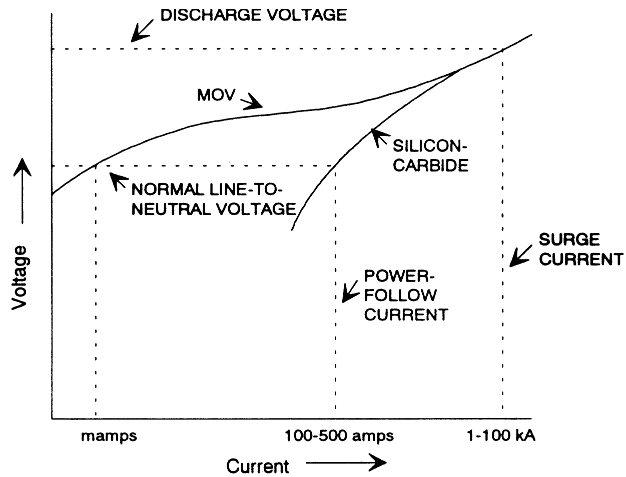


FIGURE 27-46 Illustration of power-follow current for a SiC gapped arrester and an MOV arrester.

In addition to this characteristic, it is necessary that the MOV material be capable of absorbing energy losses during conduction. The average energy absorption capability of a conventional ZnO varistor is about 150 to 200 J/cm^3. Recent advances in MOV technology under Electric Power Research Institute sponsorship resulted in new formulations and manufacturing processes for MOVs. The new MOVs are capable of absorbing 4 to 5 times more energy per volume as above.

The first designs of MOV arresters were subject to problems. The most common were thermal runaway and failures due to moisture ingress causing insulation failures. Thermal runaway can occur because an MOV is typically applied to the energized conductor without a series gap, and a small electric current always flows through the MOV during normal operating conditions. This small current and the associated ohmic losses can potentially raise the temperature of the MOV to the point that its electrical properties deteriorate, allowing more current to flow and further raising the temperature until failure. However, improvements in metal oxide materials and in packaging resulted in robust MOV arresters. The present state of the art and the superior properties of the MOV arresters make them the protection device of choice.

The first commercially available MOV arresters were gapless. As a result, the MOV block is continuously subjected to the power frequency voltage. The performance of the MOV block under this condition is very important for determining the life of the arrester. A deep understanding of these characteristics is essential for the proper application of MOV arresters.

First, observe that the voltage-current characteristic for an MOV arrester in Fig. 27-46 exhibits a knee for small currents (in the milliampere region). The knee is defined to be near the voltage required across the MOV block to cause 1 mA/cm^2 of current to flow through the block. It is important to mention that while the voltage-current characteristic of an MOV block in the protection region (high current flow) is insensitive to the temperature of the block, the voltage-current characteristic near the knee is temperature-sensitive. This is illustrated in Fig. 27-48.

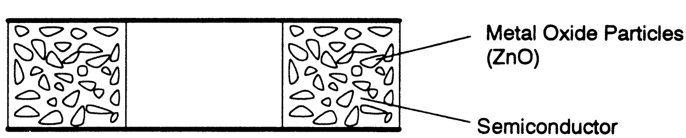


FIGURE 27-47 Structure of a metal oxide varistor (MOV). Size is exaggerated for illustration purposes.

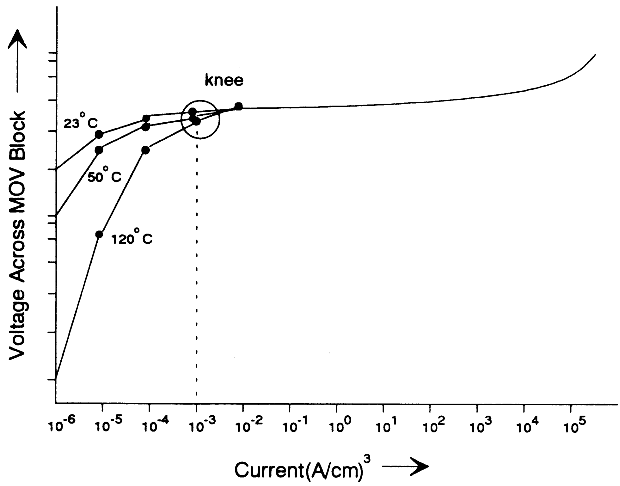


FIGURE 27-48 Voltage-current characteristic is temperature-sensitive near the knee of the curve.

Figure 27-49 illustrates the voltage versus current curves for an MOV arrester for different levels of applied voltage (sinusoidal, 60 Hz) expressed in pu of arrester rating. Note that for an applied voltage near nominal (1.05 pu), the electric current through the arrester is mostly capacitive (voltage and current are approximately 90° apart) and of low value on the order of a milliampere. As the voltage increases, the electric current increases much faster. It can be observed from the figure that the increase of the current occurs in the component, which is in phase with the voltage (resistive current) while the capacitive component remains constant.

The current flow through the arrester is responsible for power loss within the MOV block, which increases the temperature of the block. A typical power loss—applied voltage function is illustrated in Fig. 27-50. Note that as the voltage increases the power loss increases disproportionately. Increased losses and increased operating temperature have a detrimental effect on the MOV arrester life. Figure 27-51 illustrates the expected MOV arrester life as a function of continuously applied sinusoidal voltage.

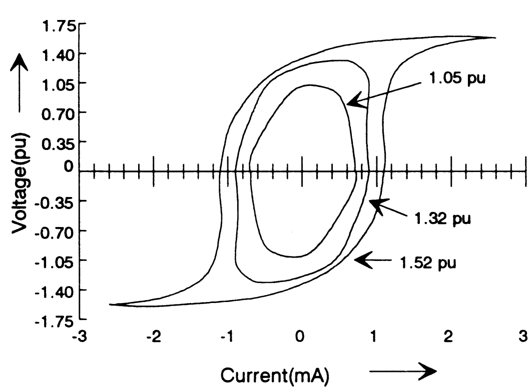


FIGURE 27-49 Illustration of voltage vs. current through an MOV for different levels of a sinusoidal (60-Hz) applied voltage.

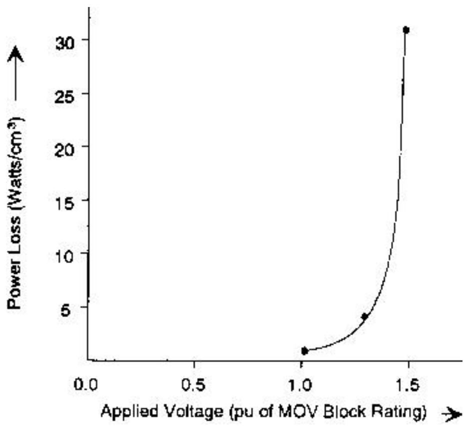


FIGURE 27-50 Typical power loss of an MOV block vs. applied voltage.

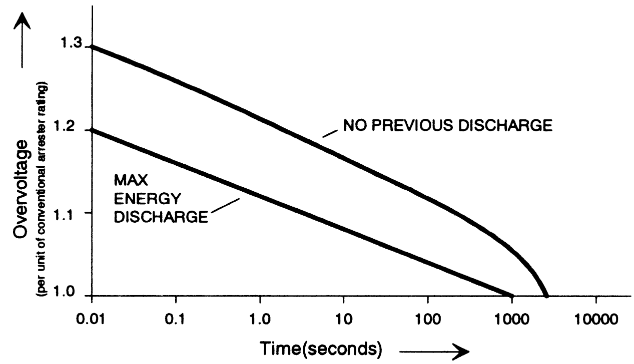


FIGURE 27-51 Expected MOV arrester life as a function of continuously applied power frequency voltage.

An MOV block, because of its dimensions, exhibits a capacitance and an inductance. Both of them are insignificant at low frequencies. However, at higher frequencies they affect the performance of the arrester. This performance is illustrated in Fig. 27-52, which shows the discharge voltage across the MOV arrester as a function of time for specific values of electric current through the arrester. From this figure it is obvious that the protective level of an MOV arrester will depend on the waveform of the overvoltage.

This behavior of the MOV surge arrester can be captured with the model of Fig. 27-53. Note that this model consists of two parallel resistance-inductance (*R-L*) circuits, a parasitic capacitance, and two nonlinear resistors. The nonlinear resistor RN_0, represents the nonlinear characteristics of the arrester to fast front end surges, while the nonlinear resistor RN_1 represents the nonlinear

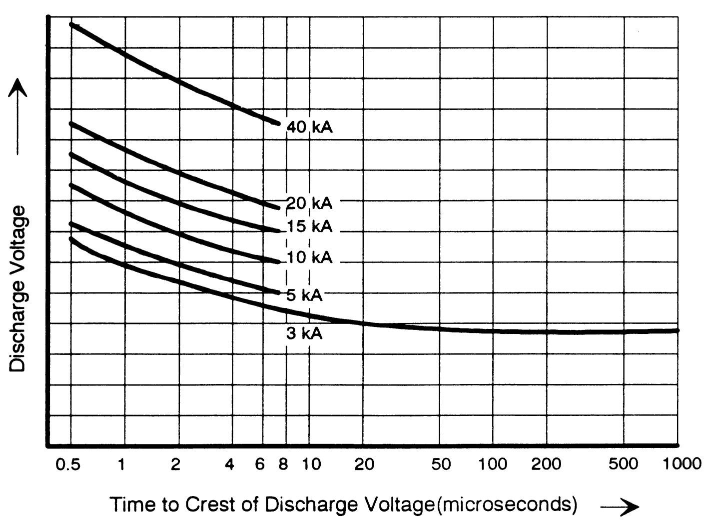


FIGURE 27-52 Discharge voltage across an MOV arrester vs. time for various current levels through the arrester.

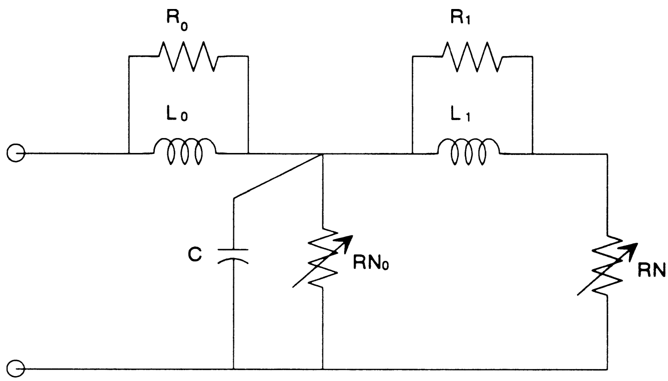


FIGURE 27-53 Nonlinear MOV arrester model.

characteristics of the arrester to slow front-end surges. The parameters of the model are selected to match the measured discharge characteristics of the arrester. Typical parameters are illustrated in Fig. 27-54.

It can be observed in Fig. 27-48 that, as the arrester discharge current increases in an MOV, the desirable nonlinear characteristics are deteriorating. This occurs in Fig. 27-48 at discharge currents above 10 kA. Specifically, the discharge voltage at currents above 10 kA increases, resulting in inferior protection characteristics. In this aspect SiC arresters are superior to MOVs because SiC arresters do not exhibit this abrupt increase of discharge voltage at higher currents. The performance of MOVs at higher discharge currents can be improved by equipping them with a shunt gap (Westrom 1990). Such a construction is illustrated in Fig. 27-55. Specifically, an MOV block in an arrester is equipped with a shunt spark gap. The gap is designed to spark over whenever the discharge current through the arrester exceeds a certain value, for example 10 kA. In this way, when the

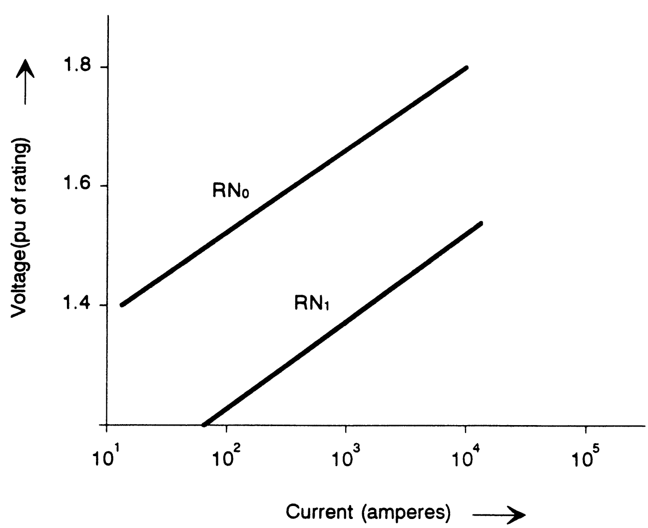


FIGURE 27-54 Typical nonlinear voltage-current characteristic of the nonlinear resistor RN_0 and RN_1.

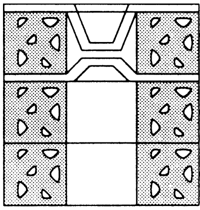


FIGURE 27-55 Typical construction of a shunt-gap MOV arrester.

discharge current increases beyond the value at which the desirable nonlinear characteristics of the MOV block start to deteriorate, the shunt gap sparks over, resulting in reduced discharge voltage and therefore improved protection characteristics.

Protective Margins. Another important concept is the protection margin (PM). The protective margin is related to the properties of the protective device relative to the withstand capability of the power apparatus under protection. For the ideal situation depicted in Fig. 27-39, the protection margin is defined as

$$\mathrm{PM} = \frac{V_w - V_p}{V_p}$$

where V_w is the withstand capability of the power apparatus under protection and V_p is the protection level.

Because both the protection level of a surge arrester and the withstand capability of a power apparatus depend on the rise time of the transient overvoltage, it is expedient to define protective margins for the standard surge waveforms. In this sense, the following protective margins are defined.

1. Equivalent front-of-wave protective margin ($\mathrm{PM}_{\mathrm{fow}}$)

$$\mathrm{PM}_{\mathrm{fow}} = \frac{V_{\mathrm{cww}} - V_{p,\mathrm{fow}}}{V_{p,\mathrm{fow}}} \times 100$$

where V_{cww} is the equipment chopped-wave withstand (cww) voltage and $V_{p,\mathrm{fow}}$ is the arrester protection level for the front of wave (fow).

2. Equipment impulse protection margin (PM_i)

$$\mathrm{PM}_i = \frac{\mathrm{BIL} - V_{p,d}}{\mathrm{BIL}} \times 100$$

where BIL is the equipment basic insulation level and $V_{p,d}$ is the arrester protection level for a full impulse wave.

3. Equipment switching protective margin (PM_s)

$$\mathrm{PM}_s = \frac{\mathrm{BSL} - V_{p,ss}}{V_{p,ss}} \times 100$$

where BSL is the equipment basic switching insulation level and $V_{p,\mathrm{ss}}$ is the arrester protective level for a switching surge.

The definition of these three protective margins is illustrated in Fig. 27-56. In this definition one should bear in mind that the arrester protective levels may be defined differently for gapped and gapless arresters. Specifically, the definition of the protective levels for gapped and gapless arresters follows.

- $V_{p,\mathrm{fow}}$, arrester protection level for front of wave

  Series-gapped arresters—front-of-wave sparkover: standard test dictates the rate of rise of front of wave to be 5.89 pu V/μs

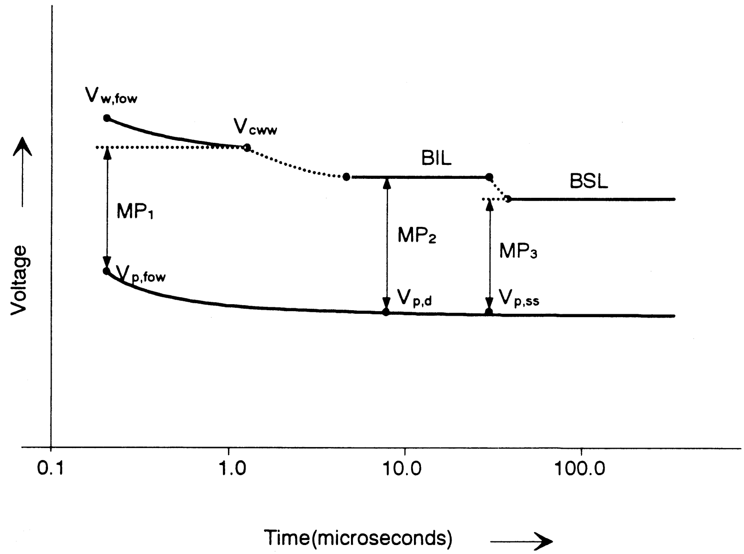


FIGURE 27-56 Illustration of the protection margin definition.

Gapless MOV arresters—front-of-wave sparkover: the highest discharge voltage to a surge-current impulse with 0.5 μs rise time and of specified crest (typically 3 to 40 kA)
Shunt-gapped MOV arresters—front-of-wave sparkover: same as for gapless MOV arresters

- $V_{p,d}$, arrester protection level for a full impulse wave

  Series-gapped arresters—let-through level: the highest full impulse expected to cause sparkover after 8 μs
  Gapless MOV arresters—discharge voltage: the highest discharge voltage to a surge-current impulse (8×20 μs) of specified crest (typically 5 to 20 kA)
  Shunt-gapped MOV arresters—discharge voltage: same as for gapless MOV arresters

- $V_{p,ss}$, arrester protection level for a switching impulse

  Series-gapped arresters—switching surge sparkover: the highest full impulse causing sparkover at a time greater than 30 μs
  Gapless MOV arresters—switching surge discharge voltage: the highest discharge voltage to a current switching impulse of specified crest (typically 3 kA)
  Gapped MOV arresters—switching surge discharge voltage: same as for gapless MOV arresters

Arrester Classification. Arresters can be classified into the following groups:

Station arresters

Intermediate arresters

Distribution arresters

Secondary, industrial, and commercial arresters

This classification is based on kilovolt ratings. Historically, station arresters provide the lowest protective levels relative to their rating (i.e., they have the highest protection quality index) and are capable of discharging the highest amount of energy. Intermediate arresters have somewhat higher protective levels relative to their rating and are capable of discharging somewhat less energy. Distribution and secondary, industrial, and commercial arresters have even higher protective levels and are capable of discharging even less energy. This historical classification and description of arresters may not be valid anymore. For example, there are MOV products that use MOV blocks for secondary arresters of the same quality as station class arresters, and therefore exhibit a very high protection quality index. It is important, however, to adopt this classification for consistency with standards.

Specific voltage ratings, protective levels, and other mechanical characteristics of complete arrester units can be found in the ANSI/IEEE standards. As an example, Table 27-6*a*, taken from ANSI/IEEE C62.1-1989, lists the standard voltage ratings for surge arresters. Table 27.6*b* lists the

TABLE 27-6*a* Surge Arrester Voltage Ratings in Kilovolts

Secondary arresters	Distribution arresters	Intermediate arresters	Station arresters
0.175			
0.650			
	1		
	3	3	3
	6	6	6
	9	9	9
	10		
	12	12	12
	15	15	15
	18		
	21	21	21
		24	24
	25		
	27		
	30	30	30
		36	36
		39	39
		48	48
		60	60
		72	72
		90	90
		96	96
		108	108
		120	120
			144
			168
			180
			192
			240
			258
			276
			294
			312
			372
			396
			420
			444
			468
			492
			540
			576
			612
			648
			684

TABLE 27-6*b* MOV Arrester Ratings in Kilovolts*

Duty-cycle voltage (kV rms)	MCOV (kv rms)	Duty-cycle voltage (kV rms)	MCOV (kV rms)
3	2.55	144	115
6	5.1	168	131
9	7.65	172	140
10	8.4	180	144
12	10.2	192	152
15	12.7	228	180
18	15.3	240	190
21	17	258	209
24	19.5	264	212
27	22	276	220
30	24.4	288	230
36	29	294	235
39	31.5	312	245
45	36.5	396	318
48	39	420	335
54	42	444	353
60	48	468	372
72	57	492	392
90	70	540	428
96	76	564	448
108	84	576	462
120	98	588	470
132	106	612	485

*For ratings not shown, consult with the manufacturer.

MOV arrester ratings taken from ANSI/IEEE std. C62.11-1999. Another example is Table 27-7, taken from ANSI/IEEE C62.2, which provides several important arrester parameters and application data for station- and intermediate-class arresters. More accurate data for specific products are available from the manufacturers.

There are a variety of commercial transient voltage suppression devices based on MOV technology. The quality of some of these products in terms of (1) their nonlinear characteristics, (2) energy absorption characteristics, and (3) current handling capability is very high. As an example, Table 27-8 illustrates the characteristics of a 480-V rated transient voltage suppression device, HA series. Figure 27-57 also illustrates the voltage-current characteristic of this device. By a simple fitting procedure, the nonlinear characteristic of this device in the range 1 to 1000 A is approximately described by Eq. (27-1), where $n \approx 24$. This is a relatively high exponent, suggesting a rather high protection quality index.

27.7 OVERVOLTAGE PROTECTION (INSULATION) COORDINATION

The objective of overvoltage protection coordination, otherwise known as insulation coordination, is to minimize the number of insulation failures and therefore the number of interruptions. This has been a fundamental task of power engineering with a profound effect on the reliability of the system. Insulation coordination methods have evolved from the ad hoc procedures of the past to the highly sophisticated computer methods of today. Yet protection schemes remain an art, and a review of the evolution of methods provides an invaluable insight. By necessity, the methods are based on incomplete data or parameters with substantial uncertainty. As a result, present models are incomplete and interpretation of results is necessary.

TABLE 27-7 Station and Intermediate Arrester Characteristics

		Protective levels, per unit crest arrester rating				Durability characteristics			
						(1)	(2)	(3)	(4)
Ratings, kV rms	Range of application nominal system voltage, kV	Front-of-wave sparkover	1.2/50-μs sparkover	Switching surge sparkover	Discharge voltage, 10 kA, 8/20-μs wave	Duty cycle initiating surge, crest amperes	Transmission line discharge, mi	High-current withstand, crest amperes	Pressure relief, rms symmetrical amperes
Station class									
3–9	2.2–12.47	2.24–4.24	1.89–3.30	Test	1.57–1.77	10,000	150	65,000	65,000 25,000
12–15	13.2–18	2.12–2.83	1.89–2.42	not	1.57–1.70	10,000	150	65,000	65,000 25,000
21–48	18–46	2.09–2.56	1.80–2.29	required	1.56–1.70	10,000	150	65,000	40,000 25,000
60–120	69–138	1.99–2.24	1.60–1.94	1.60–1.80	1.56–1.69	10,000	150	65,000	40,000 25,000
144–240	161–287	1.83–2.22	1.57–1.70	1.57–1.61	1.56–1.79	10,000	175	65,000	40,000 25,000
258–312	345	2.06–2.17	1.56–1.70	1.57–1.61	1.56–1.58	10,000	200	65,000	40,000 25,000
372 or higher	500 or higher	1.94–2.10	1.65–1.70	1.44–1.58	1.54–1.60	10,000	200	65,000	40,000 25,000
Intermediate class									
3–6	2.4–7.2	2.47–2.83	2.24–2.83	Test not	1.77–2.36	5000	100	65,000	16,100
9–48	7.2–46	2.10–2.59	1.78–2.51	required	1.77–2.19	5000	100	65,000	16,100
60–120	69–138	1.76–2.26	1.63–1.84	2.06–2.43	1.77–2.02	5000	100	65,000	16,100

TABLE 27-8 Characteristics of a Commercial MOV-Type Device, HA Series, 480 V

Model number	Size, mm	Continuous rms voltage rating, V	Continuous dc voltage rating, V	Transient energy absorption, J	Peak current 8/20 μs, A	Varistor voltage at 1-mA dc test current			Maximum clamping voltage at 200 A, 8/20 μs, V	Typical capacitance, pF
						Min., V	Normal, V	Max., V		
V481HA32	32	480	640	450	25,000	670	750	825	1290	1300

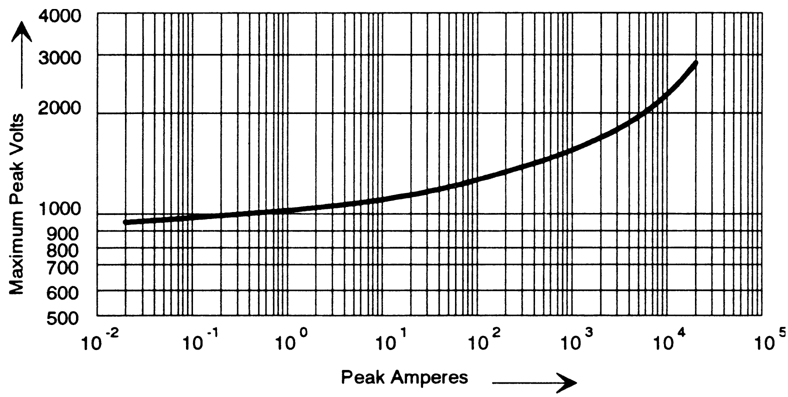


FIGURE 27-57 Voltage-current characteristic of transient voltage suppressor V481HA32, HA series.

The objectives of insulation coordination procedures also depend on the effects of insulation failure. In this respect, power apparatus may be grouped in two broad categories: (1) those that use air as insulation (external insulation) and (2) those that use solid or liquid dielectrics as insulating material (internal insulation). Insulation failure in devices of the first category is not as destructive as insulation failure in devices of the second category. With the exception of secondary effects, failure of air insulation is self-healing, since removal of the overvoltage will restore the insulation. For this reason, the objectives of insulation coordination for external insulation are typically the minimization of insulation failure. It is permissible to tolerate a small probability of insulation failure if this will minimize the cost. Probabilistic methods are pertinent for this purpose. On the other hand, the objective of insulation coordination for internal insulation is to disallow any insulation failure from any known causes. Adherence to this objective is dependent on the cost of the apparatus under protection. As an example, this objective is strictly observed for higher MVA-rating power transformers while for distribution transformers the cost of meeting the objective may not be justifiable. Because of this variability in objectives, the process of insulation coordination is examined separately for transmission lines, substations, overhead distribution systems, undergound distribution systems, and industrial/commercial systems.

Transmission Lines. Overhead transmission lines use air or porcelain as the external insulating material. Flashovers across insulator strings or phase conductor to ground wire do not lead to catastrophic failures but rather to momentary outages. In other words, the insulation strength is quickly restored after arc extinction. For this system, the objectives of insulation coordination are to minimize the momentary outages at minimum cost. This is achieved through a procedure by which the cost of momentary outages is balanced against the cost of providing for overvoltage protection. The procedure is statistical, since many driving factors are statistical, for example, lightning strength and rise time. It is illustrated in Fig. 27-58, where the distribution of overvoltages is plotted as well as the distribution of withstand

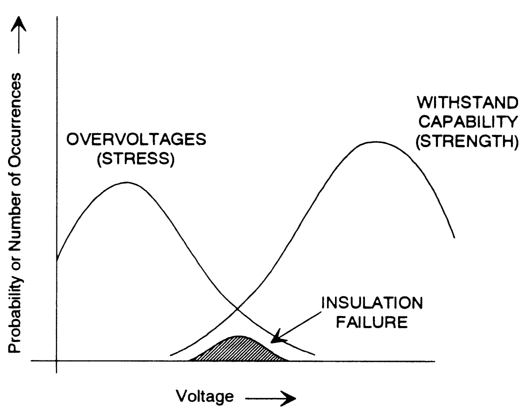


FIGURE 27-58 Statistical evaluation of insulation failures.

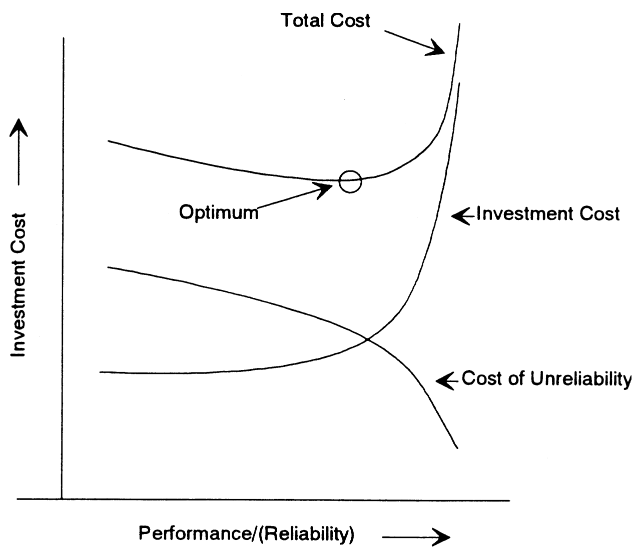


FIGURE 27-59 Optimal design procedures that provide best trade-off between performance (reliability) and cost.

capability. Outages may occur when the overvoltage exceeds the withstand capability. This procedure can be carried one step further, where the performance (reliability) of the system is balanced against cost of designs. This concept is illustrated in Fig. 27-59, where the optimum design is defined at the minimum of the total cost consisting of the investment cost plus the unreliability cost. A drawback of this method is the fact that, while the investment cost can be properly assessed, the cost of unreliability may be elusive and at best subjective. In any case, insulation coordination of overhead transmission lines involves the following basic tasks:

Design of an effective shielding against direct lightning strokes

Design of an adequate insulation strength to withstand power frequency, switching, and lightning overvoltages

Design of an effective grounding system to minimize backflashovers

The design procedure typically consists of selecting a standard design and subsequently estimating the expected number of momentary outages. Design modifications are exercised if objectives are not met. In this section, we will be concerned with estimation of the number of momentary outages. This task involves the following procedure:

Estimation of the number of lightning strokes in the area

Estimation of the percentage of strokes reaching the line

Estimation of number of shielding failures

Estimation of probability of backflashover

Effects of corona

Summary of results

Estimation of the number of lightning strokes in the area. The ground flash density is defined as the number of lightning strokes (cloud to ground) per unit of area. It is preferable to use actual data on ground flash density if available. If not, then, as a first approximation, the ground flash density is taken to be approximately proportional to the thunderstorm activity, measured in thunderstorm days.

TABLE 27-9 Empirical Relationships between Lightning Ground-Flash Density and Annual Thunderstorm Days *T*

Location	Ground-flash density, number of ground flashes/km^2 · year
India	$0.1T$
Rhodesia	$0.14T$
South Africa	$0.04T^{1.25}$
Sweden	$0.004T^2$ (approx.)
United Kingdom	aT^b ($a = 2.6 \pm 0.2 \times 10^{-3}$; $b = 1.9 \pm 0.1$)
United States (north)	$0.11T$
United States (south)	$0.17T$
United States	$0.1T$
United States	$0.15T$
U.S.S.R. (Former)	$0.036T^{1.3}$
World (temperate climate)	$0.19T$
World (temperate climate)	$0.15T$
World (temperate climate)	$0.13T$

Source: EPRI (1975).

Empirical formulas have been developed and are listed in Table 27-9. For the United States, the most commonly used formula is

$$N_1 = 0.12T \qquad \text{or} \qquad N_2 = 0.31T$$

where N_1 = ground flash density per square kilometer per year
N_2 = ground flash density per square mile per year
T = number of thunderstorm days

It must be emphasized that this formula should be viewed as an average and approximate only, since thunderstorm activity may vary from year to year and the lightning activity in a thunderstorm day may vary.

Estimation of the Percentage of Strokes Reaching the Line. The number of strokes reaching the line (shield wire or phase conductor) can be estimated with two basic models:

The geometric model

The electrogeometric model

These basic models have many variations, as many researchers have made improvements to improve their correlation to observed data.

The principle of the geometric model is illustrated in Fig. 27-60. The model postulates that the line casts a shadow around it of width W. Any ground flash which would terminate at the shadow of width W if the line was not present will be intercepted by the line. Assuming that the ground flashes are uniformly distributed over the surface of the ground, the following equation applies for the number of strokes to the line:

$$N_{t1} = 0.00012\ WT$$

or

$$N_{t2} = 0.000193\ WT \tag{27-2}$$

where N_{t1} = number of strokes to the line per kilometer of line length per year
N_{t2} = number of strokes to the line per mile of line length per year
W = width of the line shadow, m
T = number of thunderstorm days in the area

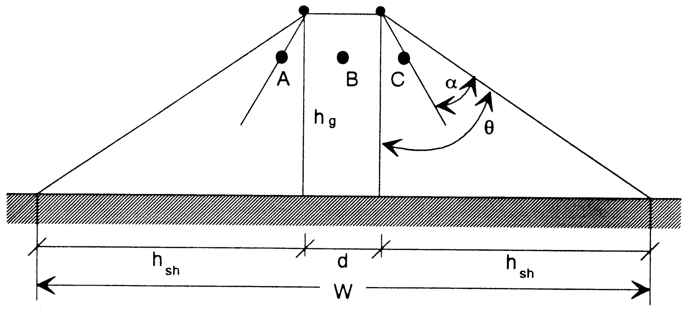


FIGURE 27-60 Illustration of application of the geometric model.

An expression for the shadow width is

$$W = 4h + d$$

where d = distance between shield wires (zero if only one shield wire)
h = effective height of the line, which can be approximated with $h = h_g + {}^2\!/_3\,(h_g - h_{gw})$
h_g = shield wire height at the tower or pole
h_{gw} = shield wire height at midspan

The geometric model is very simplistic. It fails to take into account that the striking distance of lightning is determined by the magnitude of the return stroke as well. An improved method is the so-called electrogeometric model. The principle of this model is illustrated in Fig. 27-61. This model is based on the observation that the point where the stepped leader of a lightning terminates is not determined until the stepped leader is within a striking distance. The striking distance is defined as the length of the final leg of the stepped leader which establishes contact with a ground object. The striking distance was

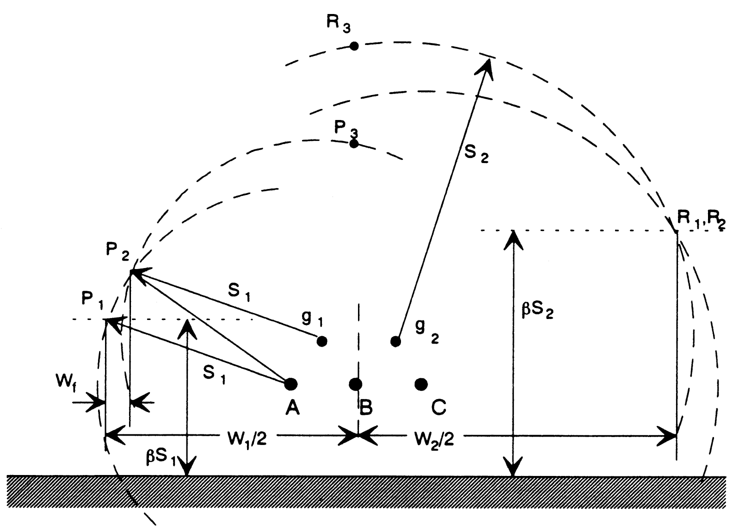


FIGURE 27-61 Illustration of application of the electrogeometric model. [*Note:* The value of β is a function of tower height. Anderson (1975) uses β = 0.67 for ultra-high-voltage (UHV) lines.]

TABLE 27-10 Proposed Equations for Determining the Striking Distance

S = striking distance in meters; I_s = first return stroke current, kA

Proposed formula	Reference
$S = 2I_s + 30\,(1 - e^{-0.147I})$	Darveniza (1979)
$S = 10\,I_S^{0.65}$	Love (1973)
$S = 9.4\,I_S^{0.67}$	Whitehead (1977)
$S = 8I_S^{0.65}$	IEEE (1987)
$S = 3.3\,I_S^{0.78}$	Suzuki (1981)
$S = 0.67\,h^{0.6}\,I^{0.74}$	Erickson (1982)
$S = 1.57\,h^{0.45}\,I^{0.69}$	Rizk (1994)

found to depend on the amplitude of the first return stroke. A number of empirical formulas have been proposed for determining the striking distance. Table 27-10 lists the most common ones.

The equation most commonly used in the United States is the second one listed in Table 27-10. Using a striking distance, which depends on stroke magnitude, generates a model, which suggests that the number of strokes intercepted by the line depends on their magnitude. Figure 27-61 illustrates the application of the method to compute the number of strokes terminating at the line for two different stroke magnitudes, 7 and 12 kA. Specifically, for strokes of 7 kA, the striking distance is S_1. For this striking distance, one can construct the width over which the strokes will be intercepted by the line. This width, W_1, is shown in the figure. It is constructed by (1) drawing a line above ground at height βS_1, where $\beta = 0.8$ (from Anderson 1968), (2) drawing a circle of radius S_1 with center at the phase conductor, and (3) drawing a circle of radius S_1 with center at the shield wire. The construction of the width W_2 for strokes of 12 kA is also shown in the figure. The computed widths W_1 and W_2 can be used in Eq. (27-2) to estimate the number of strokes to the line. It should be understood that the process is repeated for several stroke magnitudes, spanning the distribution of lightning current magnitudes; then all the results are utilized to compute the statistical distribution of the expected strokes to the line. The average value can be extracted from this distribution.

Estimation of Number of Shielding Failures. It is important to know how many of the strokes reaching the line will actually terminate on a phase conductor or on a shield wire. The consequences are obvious. A lightning stroke terminating at the phase conductor and of sufficient magnitude will most likely result in a flashover. This is obvious when one considers the magnitude of the overvoltage. When shielding failure occurs and the lighting stroke terminates on the phase conductor, it generates an overvoltage on the stricken phase which is given by

$$v_s(\mathrm{t}) = {}^1\!/_6\,(Z_g + 2Z_t)i(t)$$

and an overvoltage on the other two phases which is given by

$$v_n(\mathrm{t}) = {}^1\!/_6\,(Z_g - Z_t)i(t)$$

where $i(t)$ = lightning stroke current
Z_g = characteristic impedance for the ground mode of propagation of surges along the line
Z_t = characteristic impedance for the line mode of propagation of surges along the line
$v_s(t)$ = lightning overvoltage on the stricken phase
$v_n(t)$ = lightning overvoltage on any one of the other two phases

For a typical transmission line design

$$Z_g \approx 800\ \Omega$$

$$Z_t = 400\ \Omega$$

In this case, the overvoltage is

$$v_s(t) \approx 300i(t)$$

$$v_n(t) \approx 60i(t)$$

Obviously, a shielding failure for a lightning stroke of 20 kA will generate an overvoltage of about 6 MV, which most likely will cause a flashover. On the other hand, a lightning stroke terminating on a shield wire will result in much lower overvoltages, if the transmission line is effectively grounded. It is, of course, possible that the overvoltages resulting from a lightning stroke to the shield wire will be of such magnitude as to cause a flashover from the shield wire or tower to the phase conductor. This event is known as a *backflashover*. In any case, the line shielding must be designed in such a way that the number of times lightning reaches a phase conductor is minimized. Occurrences of such incidents are termed *shielding failure*.

The electrogeometric model provides an adequate tool to determine shielding failure. The basic premise of the method is defined in Fig. 27-61. Consider the striking distance S_1 for a lightning stroke of magnitude I_1. Consider also the line above ground at a height βS_1, the circle of radius S_1 with center at phase A, and the circle of radius S_1 with center at shield wire g_1. The intersections of these circles define the points P_1 and P_2. The electrogeometric model states that any stepped leaders, which reach the arc P_1P_2 and of magnitude less than I_1 will terminate on the phase conductor A, resulting in a shielding failure. Any stepped leader reaching the arc P_2P_3 will be attracted by the shield wire. The percentage of lighting strokes of magnitude I_1 or lower, which will reach a phase conductor, can be estimated by comparing the length of the arc P_1P_2 to the length of the arc P_2P_3. This process can be repeated for various stroke magnitudes. Subsequently, and utilizing the statistics of the stroke magnitude distribution, the probability distribution of shielding failures can be computed. The expected number of shielding failures can also be computed.

For lightning strokes of a certain magnitude, the striking distance will be such that the points P_1 and P_2 may collapse to one point. This condition is illustrated in Fig. 27-61 for a different magnitude of stroke current I_2, where the points R_1 and R_2 coincide. Any stepped leader resulting in stroke current magnitude I_2 or higher and reaching the arc R_2R_3 will terminate on the shield wire. Any other stepped leaders reaching points to the right of points R_2 will terminate to the ground. In this case, the line is shielded for lighting strokes of magnitude greater than I_2. Now assume that the line insulation can withstand the overvoltages resulting from a direct hit on a phase conductor of magnitude I_2. In this case, a shielding failure for lighting strokes of magnitudes I_2 or lower will not result in a flashover. The line is said to be *effectively shielded*.

An extension of this statement results to the concept of the critical striking distance. The critical striking distance is the striking distance for the smallest lightning stroke current magnitude, which, if it terminates on a phase conductor, will cause a flashover. The critical striking distance can be used to define the location of shield wires which will provide an effective shielding.

Estimation of Probability of Backflashover. Backflashover can occur when the transmission tower voltage at the location where the phase insulators are suspended (crossarms) becomes so high as to initiate an arc between the crossarm and the phase conductors. The driving force for initiating the arc is the voltage across the insulator which depends on many factors

Voltage at the crossarm due to the lightning stroke

Induced voltage on the phase conductor due to surges on the shield wire generated by the lightning stroke

Power frequency voltage at the phase conductor

Tower footing impedance

Traveling-wave phenomena along the tower

Lightning strokes to the tower result in the highest overvoltage at the top of the tower as compared to overvoltages from lightning strokes on the shield wire along the span. As an example, consider a lightning stroke of crest 10 kA and rise time 1 μs terminating on the shield wire of a transmission line. For simplicity, neglect the phase conductors. The system is illustrated in Fig. 27-62*a*. All tower ground resistances are 80 Ω. The parameters of the shield wire, for the example under consideration, are

$$L = 2.7\ \mu\text{H/m}$$

and

$$C = 6.58 \times 10^{-12}\ \text{F/m}$$

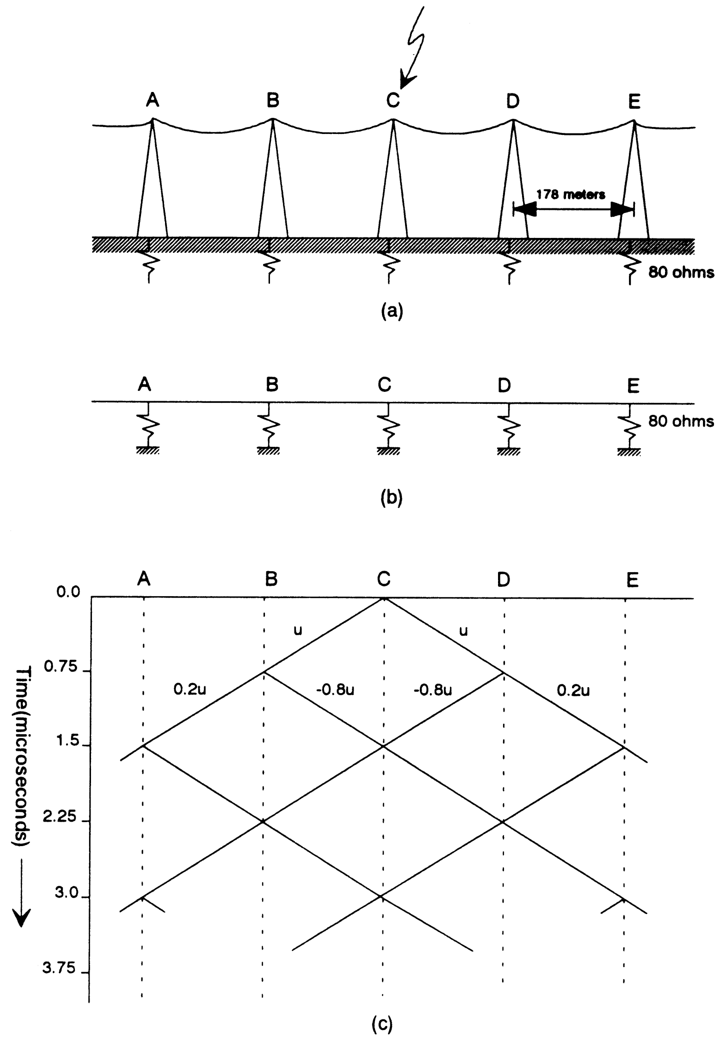


FIGURE 27-62 Simplified calculations of overvoltages resulting for a lightning stroke to a tower: (*a*) schematic; (*b*) simplified circuit; (*c*) Bewley diagram.

The characteristic impedance of the shield wire is $Z = \sqrt{L/C} = 640\ \Omega$, and the speed of propagation of surges is $v = 1\sqrt{LC} = 237$ m/μs. The travel time from one tower to the next is 0.75 μs. To further simplify the computations, traveling-wave phenomena along the tower are neglected, resulting in the simplified circuit of Fig. 27-62*b*. The reflection and transmission coefficients at the tower are $\alpha = -0.80$ and $\delta = 0.20$. The Bewley diagram for this circuit is illustrated in Fig. 27-62*c*, where $u(t)$ is computed as

$$u(t) = R_{eq} i(t) = 64 i(t)$$

Thus, $u(t)$ will have a rise time of 1 μs and crest of 640 kV. By summing up all the waves arriving at point C, the voltage waveform c1 of Fig. 27-63 is obtained. This procedure can be repeated,

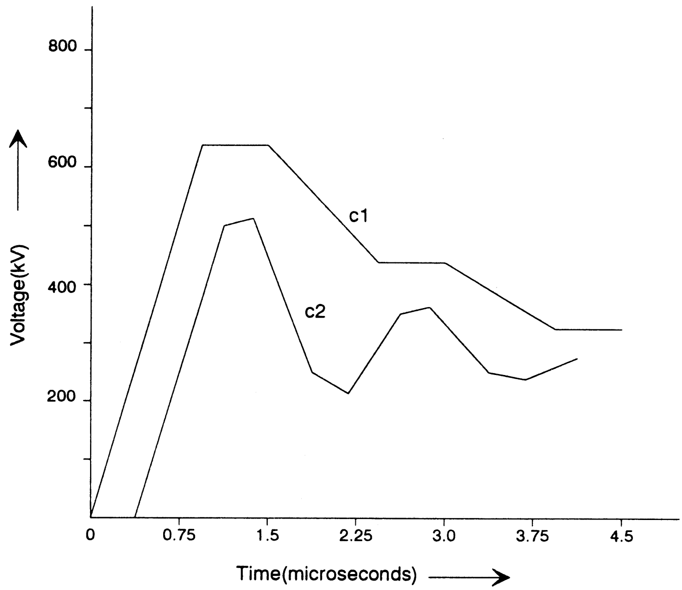


FIGURE 27-63 Lightning overvoltages at top of tower; in c1 the lightning stroke terminates at tower top, and in c2 the lightning stroke terminates at midspan.

assuming that the same lightning stroke terminates at the shield wire in the middle of the span. The resulting overvoltage at the top of the tower C is shown as waveform c2 in Fig. 27-63. Note that the crest of this waveform is much lower. In reality, the tower footing impedance is frequency-dependent and traveling-wave phenomena along the tower affect the overvoltages. Figure 27-64 illustrates a typical voltage waveform across the phase insulators using a rather sophisticated model (Cokkinides and Meliopoulos 1988). Note that reflections from adjacent towers quickly reduce the voltage across the phase insulators. Also, the crossarm voltage is in general lower than the top of the tower voltage. As a result, bottom phase insulators, in general, experience the highest

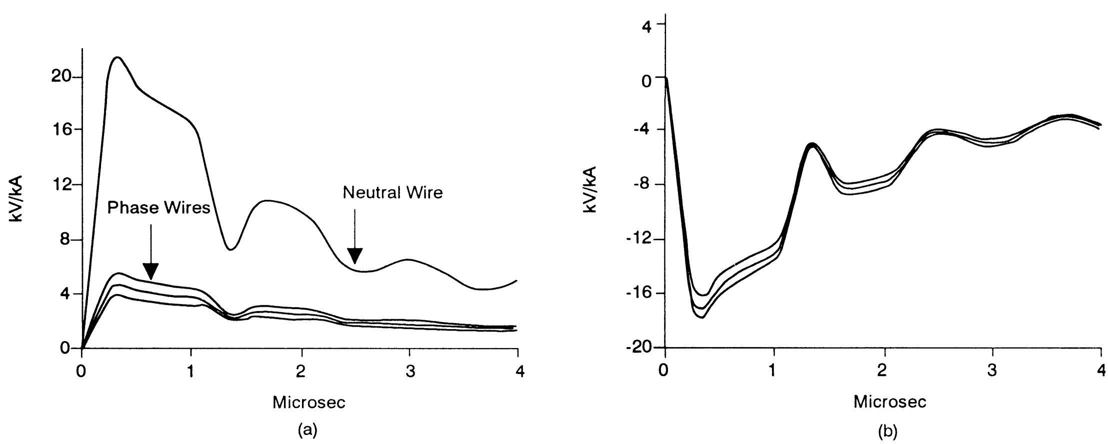


FIGURE 27-64 Typical voltage waveforms across phase insulators due to lightning: (*a*) transient voltage of phase conductors and shield wire; (*b*) transient voltage across insulators.

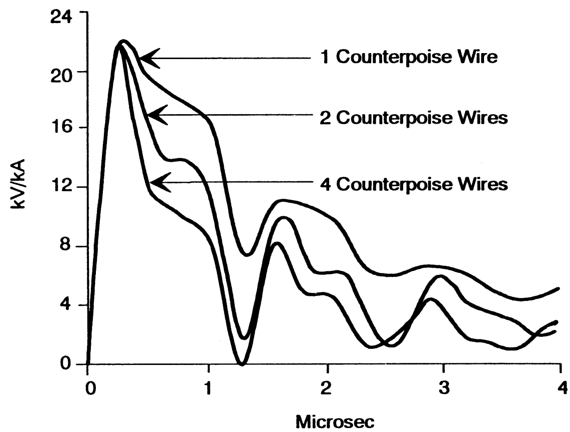


FIGURE 27-65 Illustration of effects of tower grounding on voltage at top of tower due to a lightning stroke at the tower.

overvoltage for backflashover. The tower ground also affects the overvoltage. As an example, Fig. 27-65 illustrates the effect of tower grounding on the lightning voltage at the top of the tower. Note that the voltage at the top of the tower is not affected by tower grounding for the first approximately 0.2 μs after lightning initiation. However, the overall overvoltage waveform is drastically affected by tower grounding. Since the characteristics of arc initiation across the insulator string have a volt-time characteristic, it is very important to examine the effect of voltage waveform on backflashover initiation.

CIRGE has accepted the following volt-time curve for line insulator flashover, which is consistent with the work of Darveniza et al. (1975, 1979) and Whitehead (1977)

$$V_{\bar{f}} = 0.4W + \frac{0.71W}{t^{0.75}}$$

where W = line insulator length in meters
t = time to breakdown in microseconds
$V_{\bar{f}}$ = flashover voltage in megavolts for negative surges

One should compare the volt-time characteristic to the actual voltage across the phase insulators to determine whether a backflashover will occur. Figure 27-66 illustrates such a construction where the insulator withstand capability versus time of a 1-m-long insulator string is superimposed on the actual voltage experienced by the insulator.

Effects of Corona. The overvoltage occurring at a phase conductor or shield wire establishes an extremely high-voltage gradient perpendicular to the conductor. This gradient, wherever it exceeds the withstand capability of air, will generate electric discharges, which will electrify the air surrounding the conductor. This phenomenon is known as *corona*. Corona increases the apparent radius of the conductor while it does not affect the geometric mean radius of the conductor. As a result, it decreases the characteristic impedance of the line and the level of overvoltages resulting from lightning.

The voltage gradient (electric-field intensity) around an energized conductor is approximately

$$E = \frac{V}{r \ln (2h/a)}$$

where V is the conductor voltage, r is the distance from the conductor centerline, h is the conductor height above the earth, and a is the conductor radius.

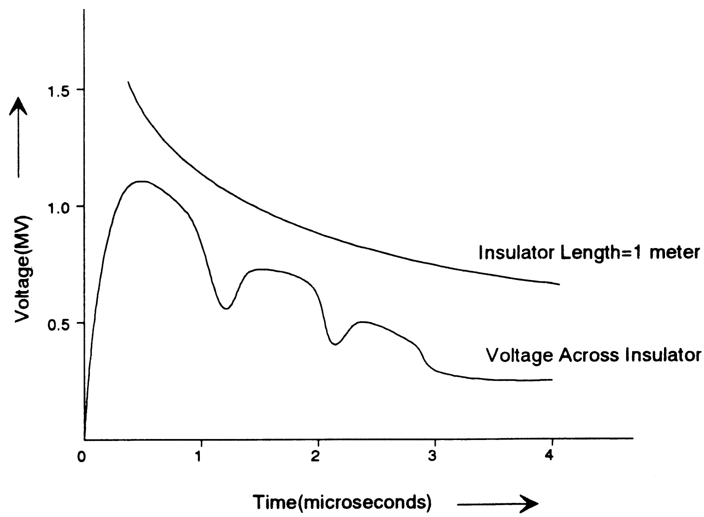


FIGURE 27-66 Insulator voltage-time curve and insulator voltage.

The value of 15 kV/cm is accepted as the electric-field intensity required to establish corona. Thus, corona will exist in a radius a_c such that

$$15 \text{ kV/cm} = \frac{V}{a_c \ln (2h/ac)}$$

As an example, a conductor 50 ft above the earth which experiences an overvoltage of 3.5 MV will establish corona at a radius $a_c = 2$ ft.

Summary of results. In summary, methods have been discussed for design of line-shielding systems, estimation of shielding failure, estimation of overvoltage stresses across insulators, and methods to determine the possibility of flashover across an insulator. There is a considerable uncertainty with almost all parameters and models affecting shielding failure, magnitude, and waveform of overvoltage across insulators, etc. This uncertainty can be dealt with by using probabilistic methods. Monte Carlo simulation is utilized for this purpose and is described in Sec. 27.8.

Substation Lightning Insulation Coordination. Substations typically comprise equipment such as transformers, bus work, breakers, reactors, and capacitor banks etc. This equipment must be protected against lightning overvoltages. As a rule, substations are shielded against direct lightning strikes on phase conductors. Thus, lightning overvoltages in a substation may occur for the following reasons:

1. Backflash due to a strike in the shield of the substation

2. Direct lightning hit due to shield failure

3. Lightning surges from lines (most common)
   a. Line shielding failure
   b. Line backflashover
   c. Induced voltage

The overvoltage protection of substations consists of shielding against direct lightning strikes and application of protective devices (surge arresters) to protect specific power apparatus.

Substation shielding methods are similar to those applied for transmission lines. One can use the geometric model or the electrogeometric model to determine the effectiveness of the shielding wires, masts, etc. The application of the electrogeometric model for substation shielding analysis is more complex since the analysis must be performed on a three-dimensional basis as opposed to a two-dimensional basis for a transmission line. A simplified variation of the electrogeometric model for the three-dimensional problems is the so-called rolling-sphere method. The basic idea of the rolling-sphere method is illustrated in Fig. 27-67*a*. Let S_c be the critical striking distance as defined for transmission lines:

$$S_c = 10 I_c^{0.65}$$

where I_c is the critical stroke current, kA, as defined for transmission lines. The rolling-sphere method postulates that by rolling a sphere of radius S_c over the shield system of the substation, the protected equipment are those which are not crossing the path of the sphere as is illustrated in Fig. 27-67. This method provides a simple procedure to design an effective shielding system for the substation. Figure 27-67*b* provides a visuailization of the application of this method to an actual substation.

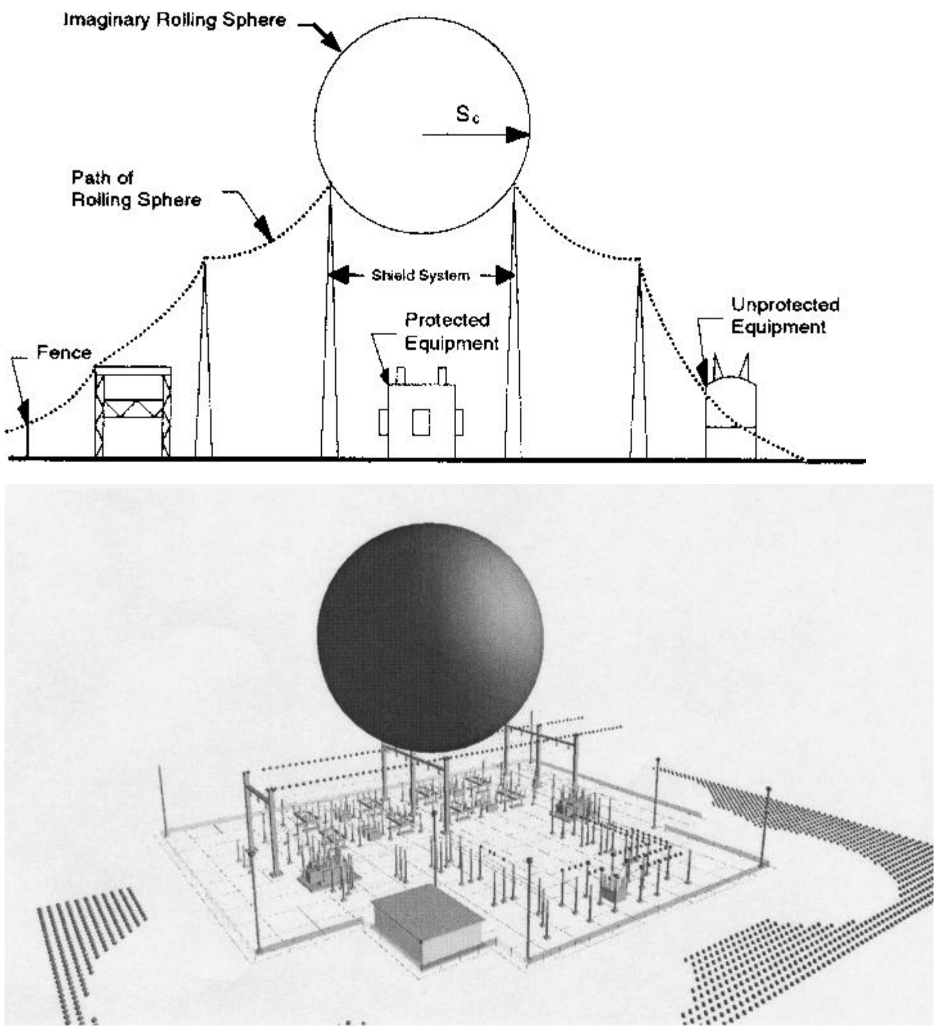


FIGURE 27-67 Illustration of the basic idea of the rolling-sphere method.

In addition to the shielding system, protection must be provided against overvoltages resulting from (1) lightning strikes to the shielding system and (2) surges entering the substation from the transmission lines (which are the most common). The magnitudes of the overvoltages in (1) above depend on the design of the grounding system. On the other hand, the overvoltages from (2) above depend on many factors, such as distance of point of initiation and source of surge (direct hit, induced). These overvoltages are typically limited by the insulation level of the transmission line. Since the usual case is that the substation equipment will have a BIL below the insulation level of the lines connected to the substation, overvoltage protection for the substation equipment must be provided.

The application of protective devices (surge arresters) for the protection of power transformers is illustrated in Fig. 27-68. The application procedure consists of the following steps:

1. Select arrester rating (preliminary).
2. Determine arrester protection levels
   $V_{p,\,\text{fow}}$ = arrester protection level for a front of wave
   $V_{p,d}$ = arrester protection level for a full-impulse wave
   $V_{p,ss}$ = arrester protection level for a switching impulse
3. Determine transformer withstand capability
   V_{foww} = front-of-wave withstand
   V_{cww} = chopped-wave withstand
   BIL = basic insulation level
   BSL = basic switching insulation level
4. Plot arrester protection levels and transformer withstand capability on a common coordinate system, as illustrated in Fig. 27-68.
5. Determine protection margins.
6. If protection margins are not higher than the minimum recommended, select the next-lower arrester rating that is compatible with the normal operating voltage of the system, and go to step 2. Otherwise, go to step 7.
7. Determine whether arrester has adequate energy absorption capability for the application.

This procedure is pertinent for application of surge arresters near the transformer. If there is a separation distance between arrester and transformer, the effects of separation must be accounted for. Specifically, separation results in deterioration of protective margins, as illustrated in Fig. 27-69. The figure shows an application of an ideal arrester 15 m away from the transformer. The ideal arrester

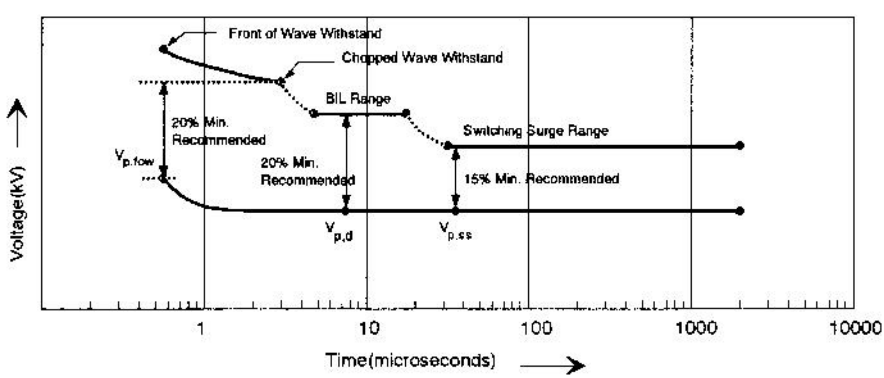


FIGURE 27-68 Application of surge arresters for protection of power transformers.

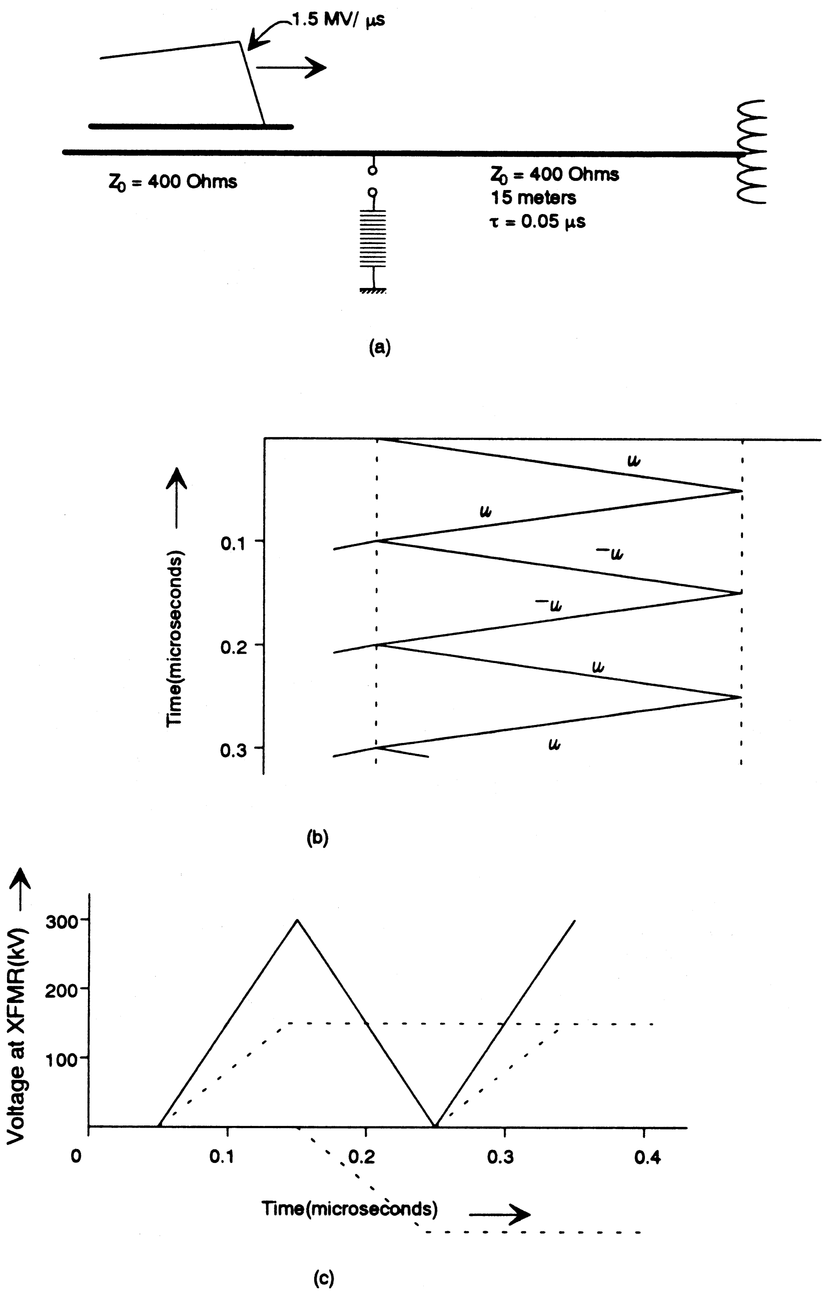


FIGURE 27-69 Example of arrester-transformer separation effects: (*a*) system configuration; (*b*) Bewley diagram; (*c*) voltage at transformer.

is rated 150 kV. Because of the separation distance, the overvoltage reaching the transformer will be 300 kV, double the protective level of the arrester.

Distribution System Overvoltage Protection. Distribution circuits are typically not insulated to withstand direct lightning strokes. As a result, direct strikes will cause a flashover. Direct strikes on distribution lines are not frequent since the poles are not as high and therefore are shielded from trees and structures. On the other hand, distribution lines may be vulnerable to overvoltages resulting from lightning strokes to nearby trees, ground, or other objects. These voltage surges are known as *induced lightning voltages* and are injected into the power system through coupling. The coupling can be conductive through the conductive soil and the power system grounding structures, inductive, or capacitive. In a typical situation, all the coupling mechanisms may be present resulting in a voltage surge to the power system. These voltages are called *induced voltage surges* and are generally much lower than those occurring after a direct strike. Specifically, they rarely exceed 500 kV. The induced lightning overvoltages are of concern for distribution lines 35 kV or below. Higher kilovolt-level lines (i.e., 69 kV and above) have sufficient insulation to withstand induced voltage surges.

The mechanism of induced voltage surges is illustrated in Fig. 27-70. A lightning stroke terminating at a location near a distribution line induces a surge on the line through conductive, inductive, and capacitive coupling. Several models to estimate the level of the induced voltage surges have been reported in the literature (Eriksson et al. 1982; Liew and Mar 1986). A simplified formula suggests the following induced voltage

$$v_p = Z_0 Ih \cdot \left[\left(\frac{1}{2y}\right) \frac{1 + (x + \beta y)}{\sqrt{x^2 + 2y^2 + 2\beta xy - \beta^2 y^2}} + \frac{2\beta x + y}{y^2 + (2\beta x + y)^2} \cdot \frac{1 + 2\beta^2 x + \beta y - \beta x}{\sqrt{x^2 + 2y^2 + 2\beta xy - \beta^2 y^2}} \right]$$

where $\beta = 0.004I^{0.64} + 0.068$ first stroke
$= 0.004I^{0.86} + 0.18$ subsequent strokes
Z_0 = distribution line characteristic impedance
I = stroke current crest

The terms x, y, and h are as illustrated in Fig. 27-70. In addition, many measurements of induced voltage surges have been performed which practically verify the available models.

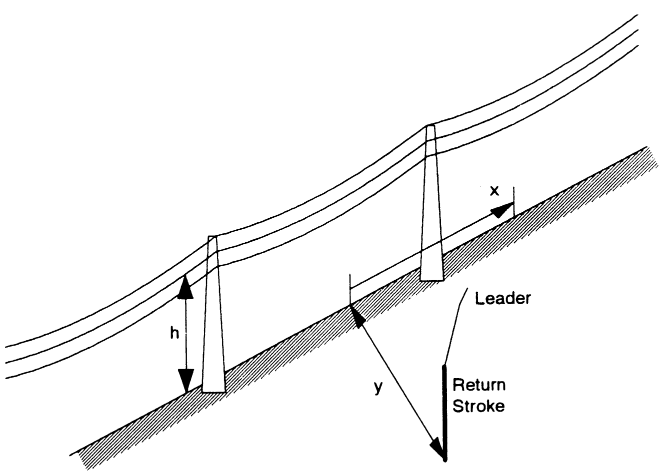


FIGURE 27-70 Illustration of lightning stroke near a distribution line.

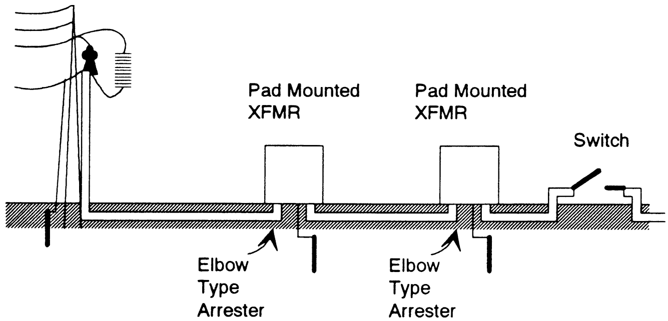


FIGURE 27-71 Typical overvoltage protection scheme for underground residential distribution (URD) systems.

Induced voltage surges on distribution lines are quite frequent. It is therefore necessary to insulate distribution lines to withstand these surges. This translates into the requirement of a 300 BIL for distribution lines. In addition, power apparatus, connected to distribution lines and with BIL lower than the induced voltage surges, such as distribution transformers (typically the BIL of distribution transformers is 100 kV) must be protected. It is practical to protect transformers with surge arresters of appropriate ratings.

Underground Distribution System Overvoltage Protection. A typical configuration of an underground distribution system is shown in Fig. 27-71. Typically a surge arrester will be applied at the riser pole. Underground residential distribution systems present a relatively low characteristic impedance (30 to 50 Ω). When a transient (surge) reaches the URD system through the high-surge-impedance overhead system, and because of the presence of a surge arrester at the riser pole, the surge transmitted to the cable will be of very fast rise time, typically a small fraction of a microsecond. The magnitude of the surge is determined by the characteristics of the arrester. This surge will propagate along the cable and will double when it arrives at an open point or at the end of the cable where a transformer may be present. If this overvoltage is below the BIL level of the URD cable, no additional protection is required. However, the BIL of typical URD cable is relatively low. As an example, Table 27-11 illustrates the typical BIL for the most usual cable classes. In most cases, if a cable is protected with a surge arrester at the riser pole only (of the recommended arrester rating), it is possible that the overvoltage at the open end of the cable exceeds the BIL of the cable with the potential of failure. Use of surge arresters at the open ends, or, better yet, use of elbow type arresters at each transformer, drastically improves cable and transformer protection.

TABLE 27-11 Typical Basic Insulation Levels for Cables

Cable kV class	Typical BIL	Recommended arrester rating
15	95	9
25	125	18
34.5	150	25

Overvoltage Protection of Industrial and Commercial Systems. Industrial and commercial power systems are subject to overvoltages resulting from lightning or switching operations. By far, lightning overvoltages are the most stressful. Since these systems are interconnected to power systems, disturbances on the power system will be transmitted to them. These systems are also subjected to direct lightning. Therefore, these systems also require shielding against lightning. Typically, a grounding system will be installed as well as a lightning protection system (shielding) to divert any direct lightning strokes to ground. This grounding system is referred to as *external grounding* to distinguish it from the so-called internal grounding, which refers to the grounding system of various equipment in the facility. The external grounding systems of industrial/commercial power systems are interconnected to the power system, as shown in Fig. 27-72. The lightning protection system is basically a shielding system designed to route the lightning surges into the external grounding system for the purpose of minimizing potential differences within the facility. A system like this is subjected to lightning overvoltages which may enter from a number of points. Table 27-12 lists a number of possible points of entry. A coordinated design of the external grounding system, lightning protection system, and internal grounding system can provide a system, which is hardened against lightning and other sources of overvoltages. The effectiveness of the system can be assessed with analysis by using the methods discussed in this section.

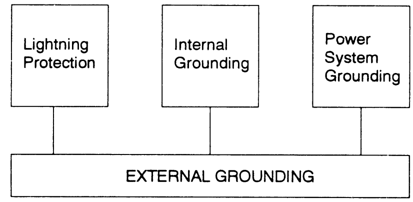


FIGURE 27-72 Conceptual description of the grounding system of an industrial or commercial facility.

TABLE 27-12 Lightning Points of Entry

Air terminals
Communication towers (if present)
Power system grounding
Fence

27.8 MONTE CARLO SIMULATION–BASED METHODS

In many parts of the globe, the performance of electrical installations is affected by lightning. In this case, it is important to design the system in such a way that outages from lightning are minimized. Typically, transmission lines, substations, and commercial and industrial buildings are the focus of the design process. Because lightning parameters present substantial variability, the effects of lightning must be evaluated on a statistical basis. Other parameters that affect system performance may also present substantial variability. One important parameter is soil resistivity, which affects ground impedances and therefore lightning overvoltages. Monte Carlo simulation methods are well suited to evaluate the effects of these parameters on system performance. A simplified description of a Monte Carlo simulation, given by Moussa and Wehling (1992), considers two parameters with substantial variability, and is illustrated in the following sequence of steps:

1. Generate a sample of lightning stroke described with respect to (*a*) crest magnitude and (*b*) rise time.
2. Generate a sample of soil resistivity.
3. Generate a sample of power frequency voltages.
4. Compute probability of shielding failure.
5. Perform an experiment and determine the lightning termination point (phase conductor, shield wire, etc.).
6. Compute transient voltages in system for conditions described above.
7. Perform effects (failure) analysis (compare voltage stresses on insulation against withstand capability).
8. If number of trials exceeded allowable, go to step 9. Otherwise go to step 1.
9. Generate histograms of maximum overvoltage and backflashover.

It should be understood that this Monte Carlo simulation requires that a good analysis procedure should be available to reliably compute the overvoltages on the insulation. Such models have been developed and are available. Detailed description of these methods is beyond the scope of this text. Instead, some typical results will be presented and discussed.

A typical application of these methods is to evaluate the performance of a specific power line. Figures 27-73 and 27-74 respectively illustrate examples of shielding analysis for two different

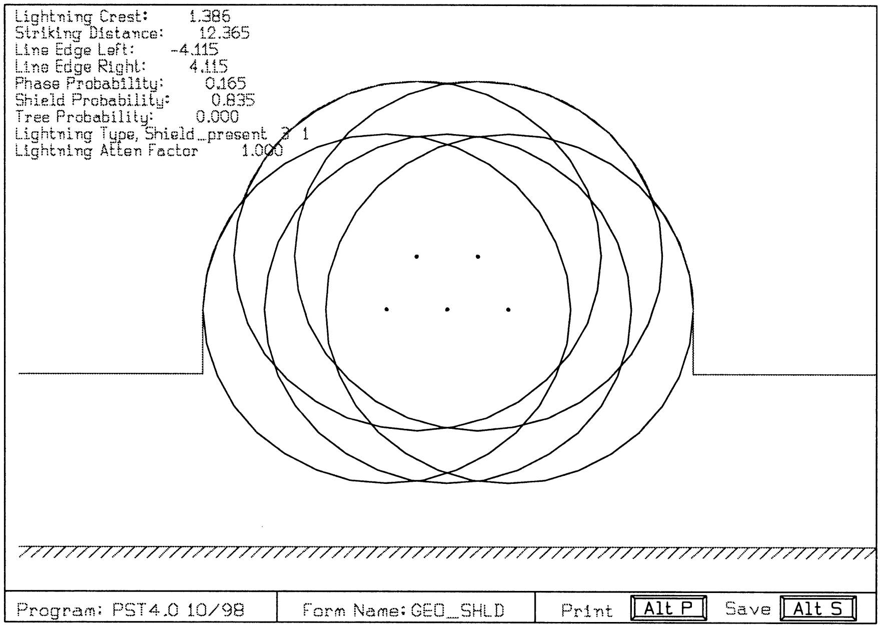


FIGURE 27-73 Example of shielding analysis. Lightning crest is 1.386 kA. There is a finite probability of direct strike on the phase conductor.

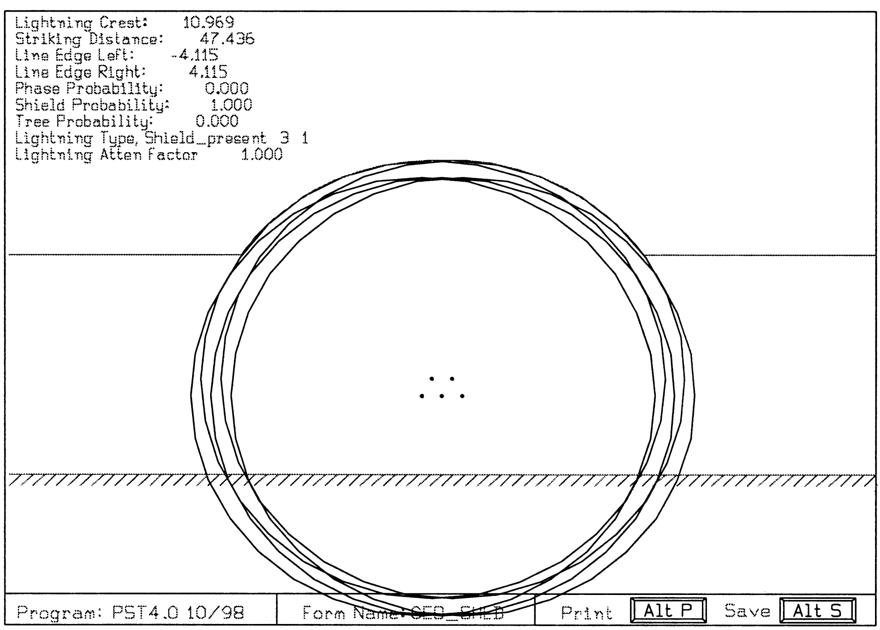


FIGURE 27-74 Example of shielding analysis. Lightning crest is 10.969 kA. The probability of direct strike on the phase conductor is zero.

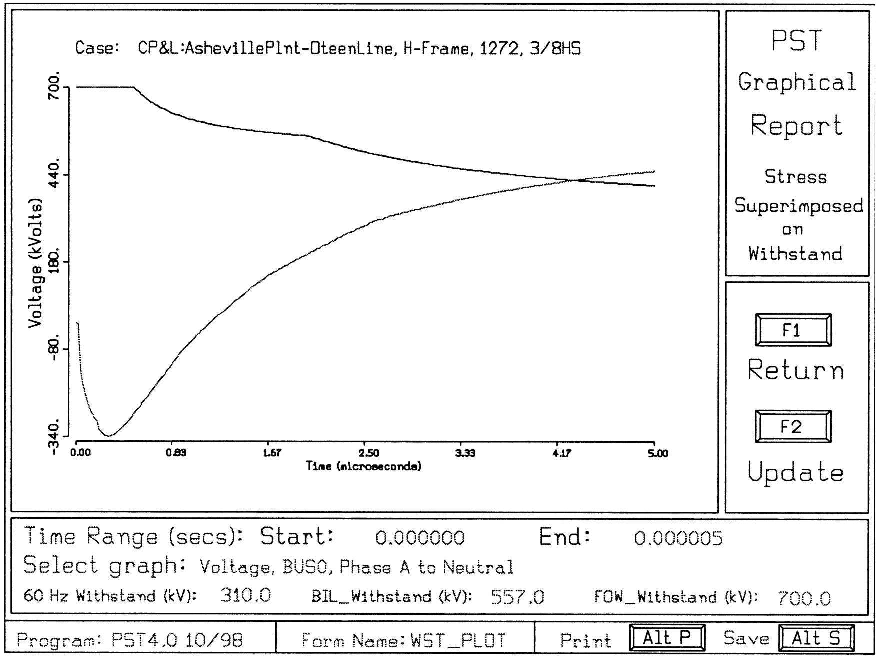


FIGURE 27-75 Example of effects analysis. Overvoltage across insulator exceeds the volt-time withstand curve of the insulator.

samples (trials) of the Monte Carlo simulation. Note that for the same system, when the current is low, there is a probability of shielding failure. As the lightning current increases, shielding failure may be eliminated.

Figures 27-75 and 27-76 respectively illustrate examples of effects (failure) analysis for two different trials of the same system. Note that the conditions illustrated in Fig. 27-75 will lead to insulation flashover, while the system will withstand the conditions illustrated in Fig. 27-76.

The results of the Monte Carlo simulation are illustrated in Fig. 27-77. Note that for this case, which refers to a transmission line, there is a small probability of direct hit to the phase conductor and that the average crest of direct hit is 8.17 kA s. The results are in terms of expected number of flashovers per 100 mi of the line and per year.

27.9 LIGHTNING ELIMINATION DEVICES

It is always desirable to eliminate the destructive effects of lightning on electrical equipment. Benjamin Franklin was the first to invent the lightning rod that provides a sacrificial path for the flow of the lightning current. Professor Moore (1997) has researched Franklin's investigations and points out that, initially, Franklin believed that his lightning rod would eliminate lightning. Experimental observations have shattered this belief and Franklin accepted the fact that the lightning rod does not eliminate lightning but rather behaves as the sacrificial electrode for the termination of lightning so that other structures in the vicinity can be protected. This observation developed into the modern shielding theory we use for design purposes and which was discussed earlier.

The desire to seek a device that prevents lightning did not die. The Czech scientist Prokop Divisch has advocated this goal in 1754. The basic idea in lightning prevention is to provide multiple points of

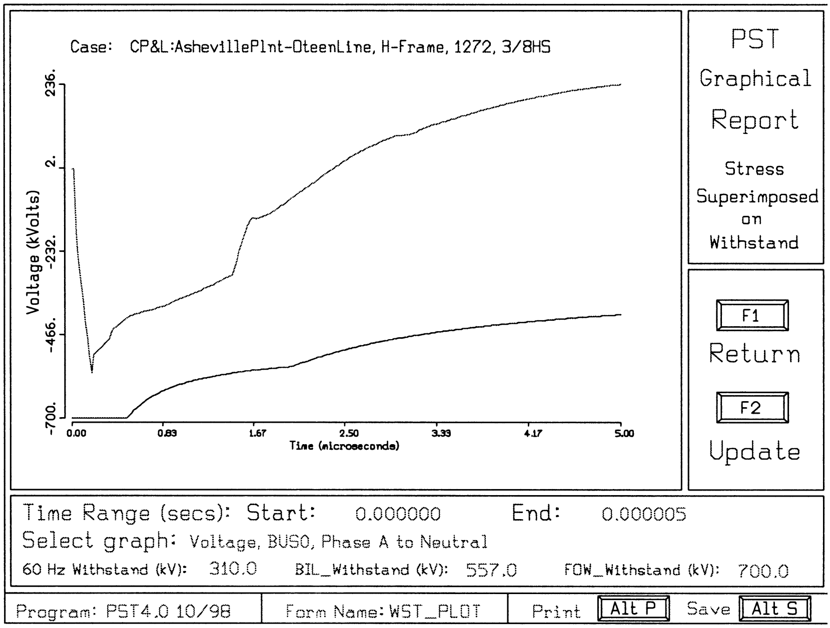


FIGURE 27-76 Example of effects analysis. Overvoltage across insulator does not exceed the volt-time withstand curve of the insulator.

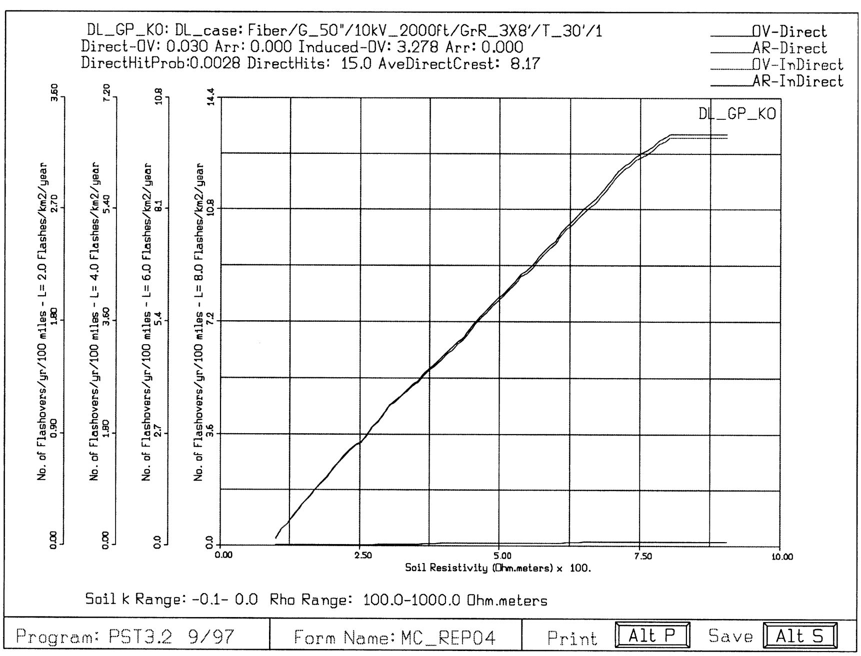


FIGURE 27-77 Example of Monte Carlo simulation results.

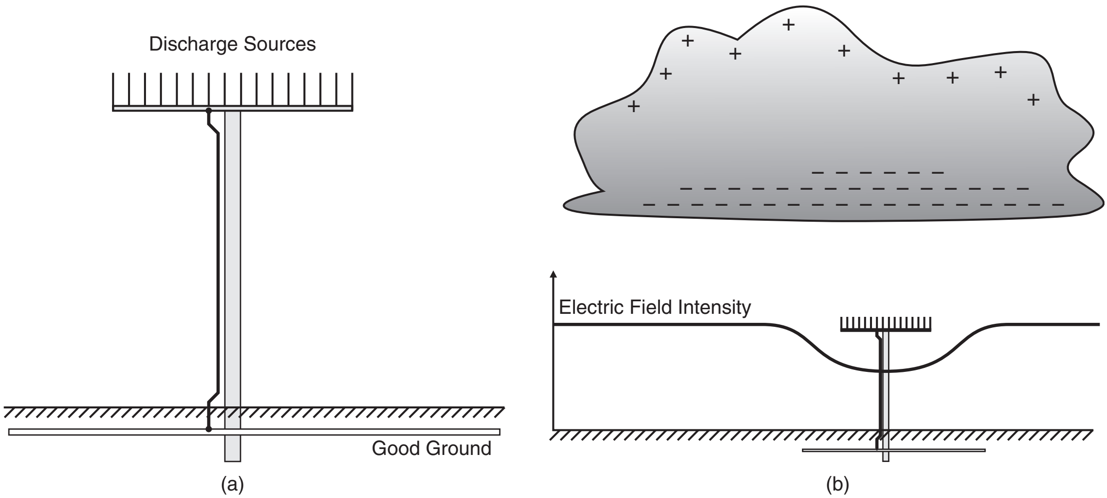


FIGURE 27-78 Conceptual illustration of a lightning elimination system showing (*a*) a good ground with discharge sources and (*b*) the discharge into the electrified cloud.

discharge to neutralize an electrified thundercloud. In theory, these charges can reach the electrified cloud and prevent the initiation of lightning. In fact, in 1930 a patent was awarded to J. M. Cage of Los Angeles for a lightning prevention device consisting of *point-bearing* wires suspended from a steel tower for the purpose of protecting petroleum storage tanks. Later, in 1971, lightning prevention systems were commercialized.

There are many technologies of lightning prevention systems: (a) active air terminal that generate ions which blend into the atmosphere to minimize the electric field around the protected facility, (b) terminals with multiple sharp edges that ionize the air around the terminal for the purpose of minimizing the electric field around the protected facility and (c) umbrella terminals, which provide a smooth surface that minimizes the maximum electric field around the protected facility. As an example the theory of operation of technology (b) is discussed. Specifically, the basic idea of this technology is illustrated in Fig. 27-78. The system consists of sharp objects (discharge sources), which are electrically connected to a good grounding system (Fig. 27-78*a*). The theory is that the discharge released by the discharge sources into the atmosphere will reach the electrified cloud (Fig. 27-78*b*) and will partially neutralize it.

The success of lightning prevention systems has not been as advertised. There is indisputable field evidence that the *lightning elimination devices* cannot eliminate lightning. On the other hand, there are cases in which these devices helped reduce the frequency and extent of damage from lightning (Moussa 1998). The present knowledge and theories of lightning are consistent with this field experience.

ACKNOWLEDGMENTS

A large number of people directly or indirectly contributed to the development of this section. Professor Nikolopoulos of the National University of Athens, Greece, introduced me to high voltage engineering principles. Dr. Roger Webb, Director of the School of Electrical Engineering, introduced and developed a course on power system transients at Georgia Tech and subsequently worked with me in teaching this course. Several folks at the Electric Power Research Institute (EPRI) supported and contributed to my work, namely, John Dunlap, Gil Addis, and Mario Rabinowitz, and numerous

EPRI project advisors contributed in the development of ideas in many different ways. Art Westrom has been an enthusiastic supporter of this work, reviewed the manuscript, and provided valuable suggestions and inputs. Finally, Fan Zhang, a Ph.D. candidate at Georgia Tech, worked with me long hours in typing the manuscript and preparing the drawings.

BIBLIOGRAPHY

Anders, G. J., and M. A. El-Kady. 1992. *Transient Ratings of Buried Power Cables*. Part I: historical perspective and mathematical model, paper 92 WM 044-8 PWRD.

Anderson, J. G. 1968. *EHV Transmission Line Reference Book*. Washington, DC: Edison Electric Institute.

Anderson, J. G. 1975. Lightning Performance of Transmission Lines, in *Transmission Line Reference Book, 345 KV and Above*. Palo Alto, Calif.: Electric Power Research Institute (EPRI).

ANSI Standard C62.2. *Guide for Application of Valve-Type Lightning Arresters for Alternating Current Systems*.

Barker, P. P., et al. 1992. Characteristics of Lightning Surges Measured at Metal Oxide Distribution Arresters, paper 92 WM 255-0 PWRD.

Berger, K. 1967. Novel Observations on Lighting Discharges: Results of Research on Mount San Salvatore. J. Franklin Institute 283(6):478–524.

Boyd, E. L. 1992. *Internal Spark Gap Protection of Distribution Transformers from Low-Side Current Surges*. IEEE Trans. Power Delivery, pp. 1592–1600.

Braunstein, A. 1970. Lightning Strokes to Power Transmission Lines and the Shielding Effect of Ground Wires. *IEEE Trans*. PAS-89(8):1900–1910.

Brown, G. W. 1978. Lightning Performance. Updating Backflash Calculation. *IEEE Trans*. PAS-97:33–38.

Chalmers, J. A. 1967. *Atmospheric Electricity*. New York: Pergamon Press.

Chowdhuri, P., and E. T. B. Gross, 1969. Voltages Induced on Overhead Multiconductor Lines by Lightning Strokes. IEEE Proc. 116:561–564.

Cokkinides, G. J. and A. P. Meliopoulos. 1988. Transmission Line Modeling with Explicit Grounding Representation. Electric Power Syst. Research 14(2):109–119.

Darveniza, M., et al. Nov./Dec. 1979. Modeling for Lightning Performance Calculations. IEEE Trans. Power *Appar. Syst*. PAS-98(6):1900–1908.

Darveniza, M., F. Popolansky, and E. R. Whitehead. July 1975. Lightning Protection of UHV Transmission Lines. Electra (41): 36–39.

Dean, D. S. 1966. Insulation Tests for the Design and Uprating of Wood-pole Transmission Lines. IEEE Trans. PAS-85:1258.

Dommel, H. W. 1969. Digital Computer Solution of Electromagnetic Transients in Single and Multiphase Networks. IEEE Trans. PAS-88:388–399.

Dommel, H. W., and W. S., Meyer, 1974. Computation of Electromagnetic Transients. Proc. IEEE 62:983–993.

Durie, R. C., and C. Pottle, 1992. An Extensible Real-Time Digital Transient Network Analyzer, paper 92 WM 255-0 PWRD.

EPRI. 1984. Report EL-3608, *Single-Phase, Multiphase, and Multiple-Circuit Lightning Flashovers of Transmission Lines*. Palo Alto, Calif.: EPRI.

EPRI. 1990. Report EL-6782, *Characteristics of Lightning Surges on Distribution Lines*. Palo Alto, Calif.: EPRI.

Eriksson, A. J., M. F. Stringefellow, and D. V. Meal 1982. Lightning Induced Overvoltages on Overhead Distribution Lines. IEEE Trans. PAS-101:960–968.

Fink, D. G., and H. Wayne Beaty. 1993. *Standard Handbook for Electrical Engineers*, 13th ed. New York: McGraw-Hill.

Flugum, R. W. 1970. Operation of Lightning Arresters on Abnormal Power Frequency Voltages. IEEE Trans. PAS-89(7):1444–1451.

Furukawa. S. et al. 1989. Development and Application of Lightning Arresters for Transmission Lines. IEEE Trans. Power Delivery 4(4).

Gallagher, T. J., and A. J. Pearmain. 1983. *High Voltage Measurements, Testing and Design*. New York: Wiley.

Goedde, G. L., R. C. Dugan, and L. D. Row. 1991. Full Scale Lightning Surge Tests of Distribution Transformers and Secondary Systems. IEEE Proc. Dallas, Tex. PES T&D Conference pp. 577–581.

Golde, R. H., ed. 1977. *Lightning*. New York: Academic Press.

Hileman, A. R., P. R. Leblanc, and G. W. Brown. 1970. Estimating the Switching Surge Performance of Transmission Lines. IEEE Trans. PAS-89:1455–1466.

IEEE. 1987. P*ower System Transient Recovery Voltages. Seminar Report 87THO176-8-PWR*. New York: IEEE.

Ishida, K., et al. 1992. Development of a 500 kV Transmission Line Arrester and Its Characteristics. IEEE Trans. Power Delivery 7(3):1265–1274.

Johnson, I. B., et al. 1979. *Surge Protection in Power System*. IEEE Tutorial Text 79 EH0144-6 PWR.

Kolcio, N., et al. 1992. Transient Overvoltages and Overcurrents on 12.47 kV Distribution Lines; Computer Modeling Results, paper 92 WM 273-3 PWRD.

Kroninger, H. 1974. *Further Analysis of Prof. Berger's San Salvatore Lightning Current Data*, CSIR Spec. Report ELEK 53. Pretoria, South Africa: National Electrical Engineering Research Institute.

Larson, A. 1905. *Annual Report*. Washington, DC: Smithsonian Institute, pp. 119.

Liew, A., and M. Darveniza. 1974. Dynamic Model of Impulse Characteristics of Concentrated Earths. Proc. IEEE 121(2):123–135.

Liew, A. C., and S. C. Mar. 1986. Extension of Chowdhuyri-Gross Model for Lightning Induced Voltage on Overhead Lines. *IEEE Trans*. *Power Delivery* PWRD-1(4):1073–1081.

Love, R. R., 1973. Improvements on Lightning Stroke Modeling and Applications to the Design of EHV and UHV Transmission Lines, M.Sc. thesis, University of Colorado.

McComb, T. R., et al. 1992. Digital Techniques in HV Tests: Summary of 1989 Panel Session, paper 92 WM 175-0 PWRS.

Mak, S. T. 1992. Propagation of Transients in a Distribution Network, paper 2/92 WM 272-5 PWRD.

Mancao, R. T., J. J. Burke, and A. Myers 1992. The Effect of Distribution System Grounding on MOV Selection, paper 92 WM 223-7 PWRD.

Meliopoulos, A. P. 1972. Problems and Concise Theory of High Voltage Structures (in Greek). Athens, Greece: S. B. H. Sellountos Publishing Co.

Meliopoulos, A. P. 1988. *Power System Grounding and Transients: An Introduction*. New York: Marcel Dekker.

Meliopoulos, A. P., and M. G. Moharam. 1983. Transient Analysis of Grounding Systems. IEEE Trans. PAS-102(2):389–397.

Meliopoulos, A. P., Sakis , W. Adams, and R. Casey. May 1997. An Integrated Backflashover Model for Insulation Coordination of Overhead Transmission Lines. *Electrical Power Energy Syst*. 19(4):229–234.

Melvold, D. J., S. A. Miske, Jr., and E. C. Sakshung May 15, 1974. Switching Cables Can Stress Arresters. New York: *Electrical World* magazine.

Moore, C. B. Jan. 1997. Study of Behavior of Sharp and Blunt Lightning Rods in Strong Electric Fields: Review of Lightning Protection Technology for Tall Structures. Publication AD-A075 449, pp. 96–107.

Mousa, A. M. Oct. 1998. The Applicability of Lightning Elimination Devices to Substation and Power Lines. *IEEE Trans*. *Power Delivery* 13(4):1120–1127.

Mousa, A. M., and R. J. Wehling 1992. A Survey of Industry Practices Regarding Shielding of Substation against Direct Lightning Strokes, paper 92 WM 224-6 PWRD.

Ogbuobiri, E. C., W. F. Tinney, and J. W. Walker 1970. Sparsity-Directed Decomposition for Gaussian Elimination on Matrices. *IEEE Trans. Power Appar. Syst.* PAS-89:141–150.

Papalexopoulos, A. D., and A. P. Meliopoulos 1987. Frequency Dependent Characteristics of Grounding Systems. *IEEE Trans. Power Delivery* PWRD-2(4):1073–1081.

Powell R. W. 1967. Lightning Protection of Underground Residential Distribution Circuits. IEEE Trans. PAS-86:1052–1056.

Regaller, K. ed. 1980. *Surges in High-Voltage Networks*. New York: Plenum Press.

Rizk, F. A. M. 1994. *Modeling of Lightning Incidence to Tall Structures, Part II: Application*, IEEE Transactions on Power Delivery, pp. 172–193.

Sakshung, E. C., T. J. Carpenter, and E. W. Stetson. 1965. Duty Cycle Testing of Current-Limiting Station and Intermediate Lightning Arresters. *IEEE Trans*. PAS-84(3):422–425.

Sakshaug, E. C., J. S. Kresge, and S. A. Miske, Jr. 1977. A New Concept in Station Arrester Design. *IEEE Trans*. PAS-96(2):647–656.

Sargent, M. A. 1972. Monte Carlo Simulation of the Lightning Performance of Overhead Shielding Networks of High Voltage Stations. IEEE Trans. PAS(pt. III)-91:1651–1656.

Sargent, M. A., and M. Darveniza 1969. Tower Surge Impedance. *IEEE Trans*. PAS-88:680–687.

Semlyen, A., and M. R. Iravani 1992. Frequency Domain Modeling of External Systems in an Electromagnetic Transient Program, paper 92 WM 304-6 PWRS.

Shindo, T., et al. 1992. Model Experiments of Laser-Triggered Lightning, paper 92 WM 25804 PWRD.

Smith, S. B., and R. B. Standler 1991. The Effects of Surges on Electronic Appliances, paper 91 SM 384-8 PWRD, IEEE summer meeting.

Stone, G. C., R. G. van Heeswijk, and R. Bartnikas Investigation of the Effect of Repetitive Voltage Surges on Epoxy Insulation, paper 92 WM 067-9 EC.

Suzuki, T., K. Miyake, and T. Shindo 1981. *Discharge Path Model in Model Test of Lightning Strokes to Tall Mast* IEEE Transactions on Power Apparatus and Systems, vol. PAS-100, No. 7, pp. 3553–3562.

Thomas, H. 1968. Transport Time-Delay Simulation for Transmission Line Representation. *IEEE Trans. Computers* C-17:204–214.

Tominager, S. 1980. Stability and Long-Term Degradation of Metal Oxide Surge Arresters. *IEEE Trans*. PAS-99(4):1548–1556.

Transmission Line Reference Book, 345 kV and Above. Palo Alto, Calif.: EPRI.

Udo, T. 1965. Switching Surge and Impulse Sparkover Characteristics of Large Gap Spacing and Long Insulator Strings. IEEE Trans. PAS-84:304.

Uman, M. A. 1969. *Lightning*. New York: McGraw-Hill.

Uman, M. A. 1987. *The Lightning Discharge*. New York: Academic Press.

Walsh, G. W. 1977. A New Technology Station Class Arrester for Industrial and Commercial Power Systems, Conf. Rec. IEEE Industrial and Commercial Power System Technical Conference, catalog no. 77 CH1198-11A.

Westrom, A. C. 1990. Surge Arrester with Shunt Gap. U.S. Patent 4,908,730.

Whitehead, E. R. 1977. Protection of Transmission Lines, in *Lightning*, vol. 2, R. H. Golde, ed. New York: Academic Press, pp. 697–745.

Williams, E. R. Nov. 1988. The Electrification of Thunderstorms. Scientific American pp. 88–99.

SECTION 28
STANDARDS IN ELECTROTECHNOLOGY, TELECOMMUNICATIONS, AND INFORMATION TECHNOLOGY

Marco W. Migliaro
PE, Fellow IEEE, President and CEO-IEEE Industry Standards and Technology Organization (IEEE-ISTO)

CONTENTS

28.1 INTRODUCTION

Standards are documents on which agreement has been reached, normally by consensus, that contain specifications or criteria to be used to ensure that materials, products, processes, tests, or services are suitable for their intended purpose.

Standards apply to virtually everything in the world today. The average person is not even aware of their existence, but life would not be the same without them. Engineers, computer scientists, and other scientists, however, are acutely aware of standards and their impact on the work they perform. Development of many of the original standards associated with electrotechnology was a slow process. The products, tests, or specifications being standardized were often in use in industry and had, in reality, become de facto standards before the standards that referred to them were written and approved. In general, most of the standards written were not mandatory, and it was voluntary for users to apply them. In some instances, these voluntary standards became part of government regulations, were adopted by government agencies, or were mandated by companies in specifications. When these types of events occurred, compliance with a specific standard became mandatory. Early standards also became regional to varying degrees. For example, the United States had its electric power standards and European countries had their own electric power standards. Although the two sets of standards had many similarities, there were significant differences between them. Overall, although there were some complaints, the standards development processes remained slow and the differences between standards continued to exist.

The 1980s brought some dramatic changes. The deregulation of the telephone industry, the privatization of government-run electric and telecommunication organizations, the birth of the information age, and the realization of a global market created a huge demand for standards in the fields of telecommunications and information technology. The Internet has, by itself, created the need for standards that govern its use, domain registration, and so forth. Many new players have entered the standards arena and found the traditional methods for developing standards unacceptable, primarily because they were too slow and bureaucratic. New standards in telecommunications and information technology were needed immediately, and the affected industries were moving so fast that many of the standards developed using traditional methods were obsolete before they were issued. Standards developers responded by streamlining processes and adopting fast-track systems. Even these proved too slow for some, and standards saw the birth of consortia whose charters included standards development. The global economy also caused people to take a hard look at existing regional standards. The differences between the standards were viewed as potential "barriers to trade," and harmonization efforts began to make these standards more widely accepted. For example, the United States and Canada are now attempting to harmonize their electrical wiring codes. The global market brought about another significant change in standards. In addition to the traditional standards that dealt with products, tests, and specifications, the protection of intellectual property (e.g., trademarks, inventions, and copyrights) became an important issue. Many standards developers have already begun to make their standards available on the Internet for a fee, although grassroots initiatives have continued to allow free access to standards via the Internet.

Many consortia make their standards and specifications available at no charge. A few standards development organizations have also initiated programs that allow access to a limited number of standards without fee (e.g., the IEEE's "Get 802" program).

28.2 HISTORY OF ELECTRICAL STANDARDS

Early History. The early history of electrical standards stems from activities dominated by the American Institute of Electrical Engineers (AIEE).* In 1884, the institute began actively to develop standard specifications for the growing electrical industry. In 1890, it proposed that the practical unit of self-induction be named the henry. At the same time, the institute appointed its first committee on standardization—the Committee on Units and Standards. The members of this committee were A. E. Kennelly, chairman, F. B. Crocker, W. E. Geyer, G. A. Hamilton, and G. B. Prescott, Jr. The institute also appointed a "Standard Wiring Table Committee" under the chairmanship of E. B. Crocker, to assign linear resistance of standard-conductivity copper wire and at standard temperatures.

A committee was also appointed to prepare a program for the delegates to the International Electrical Congress, held in Chicago in 1893, in regard to units, standards, and nomenclature. As a result of the congress, there were adopted units for magnetomotive force (gilbert), flux (weber), reluctance (oersted), and flux density (gauss). Subsequently, as a result of correspondence with engineering organizations in England, France, and Germany, the term *inductance* was adopted to represent the coefficient of induction (with the symbol *L*) and the present definition of the term *reactance* was proposed by Steinmetz and adopted.

First Electrical Standards. In 1896, a "National Conference of Standard Electrical Rules" was held. The institute's delegate, Professor F. B. Crocker, was made its president, and in cooperation with other national organizations, the conference promulgated the "Underwriters' Rules," which finally resulted in the *National Electrical Code*® (NEC®).**

*In 1963, the AIEE merged with the Institute of Radio Engineers to form the IEEE.
***National Electrical Code* and NEC are registered trademarks of the *National Fire Protection Association (NFPA)*.

In 1897, the Units and Standards Committee recommended adoption of the standard of luminous intensity, or candlepower, as the output of the amylacetate Hefner-Alteneck lamp. It also recommended that the Lummer-Brodhun photometer screen be adopted for measuring the mean horizontal intensity of incandescent lamps.

At the beginning of 1898, a discussion was organized on the subject of "standardization of generators, motors, and transformers." This resulted in the formation of the first AIEE product standards committee, which in 1899 published the first electrical standard under the unique title *Report of the Committee on Standardization*.

National Institute of Standards and Technology. The AIEE was a prime mover in the endorsement of a bill before the U.S. Congress, in 1901, for establishing a national standardizing bureau in Washington, DC, "for the construction, custody, and comparison of standards used in scientific and technical work." This bureau became known as the National Bureau of Standards (NBS) and has had a marked influence on the growth of U.S. technology. In 1988, the mission of the NBS was broadened by The Omnibus Trade and Competitiveness Act and other legislation, to help enhance competitiveness of U.S. industry and speed up the commercialization of new technology. At that time, the NBS was renamed the National Institute of Standards and Technology (NIST).

International Electrical Standards. In 1904, an International Electrical Congress was held in St. Louis which set a precedent for international congresses related to electrical units and standards. The congress unanimously recommended the establishment of two committees. Committee 1 consisted of government representatives and was responsible for legal maintenance of units and standards. This committee has now evolved into the International Conference on Weights and Measures (GPMU). Committee 2, of which Lord Kelvin was elected president, was responsible for standards related to commercial products in the electrical industry and became the International Electrotechnical Commission (IEC) in 1906.

Another international body, the International Commission on Illumination [Commission International de l'Eclairage (CIE)], had its first meeting in 1913. The CIE establishes international units, standards, and nomenclature, in the science and technology of light and illumination.

International Telecommunications Standards. In 1865 the first International Telegraph Convention was signed by 20 countries. This marked the formation of the International Telegraph Union (ITU). After the invention of the telephone in 1876 and wireless telegraphy (the first type of radiocommunication) in 1896, the scope of the ITU was broadened to include these new technologies. In 1906, the first International Radiotelegraph Convention was signed. The International Telephone Consultative Committee (CCIF) was formed in 1924, and the International Telegraph Consultative Committee (CCIT) was formed in 1925. In 1920 sound broadcasting began, and in 1927 the International Radio Consultative Committee (CCIR) was formed. At the Madrid Conference in 1932, the previous conventions were combined into the International Telecommunication Convention. The ITU changed its name in 1934, to the International Telecommunication Union. After World War II, the ITU became a specialized agency of the United Nations in October 1947. In 1956, the CCIF and CCIT merged to form the International Telephone and Telegraph Consultative Committee (CCITT). The year 1963 saw the first telecommunications satellite and the ITU set up a study group on space communications. Most recently, the Plenipotentiary Conference held in 1992 has remodeled the ITU to meet the challenges of the future.

International Standards Outside the Electrical Field. The original standards work in the fields outside electrotechnology was performed under the International Federation of the National Standardizing Associations (ISA), which was formed in 1926. ISA's activities ended in 1942 as a result of World War II. In 1947, the International Organization for Standardization (ISO) was established, as a result of a meeting of delegates from 25 countries that was held in London in 1946. Like the IEC, the ISO is a nongovernmental organization that promotes the development of international standardization and related activities. Its areas of responsibility are fields outside electrotechnology, light, and telecommunications.

International Information Technology (IT) Standards. When the need for international standards in the field of information technology arose, it was clear that both the IEC and the ISO needed to be involved. In 1987, an agreement between the IEC and the ISO created the Joint Technical Committee on Information Technology (JTC-1). The ITU provides input to JTC-1 as an official liasion.

Another organization, The Internet Society (ISOC), was formed in 1992. Its formation came as a result of the INET Conference held in Copenhagen in 1991, where it was decided that a neutral and internationally recognized body devoted to the support of Internet administrative infrastructure was needed.

National Standardization. Although an international standards organization for electrotechnology existed, representation on the IEC was by national committees from its member countries. Many of these countries had their own national standards organizations responsible for their national standards program, endorsement of national standards, participation in international standards development, and so forth. A number of these national organizations later became founding members of the ISO.

In the United States, five professional engineering societies and three government agencies spearheaded by the AIEE organized the American Engineering Standards Committee (AESC) in 1918. The AESC has been aptly described as a "national clearinghouse for industrial standardization." In its early years, this body was organized with 12 divisions, each based on its own area of technology. Few of these became active. The electrical engineering division actually became the strongest, even to the point of having its own bylaws. Today, the AESC is known as the American National Standards Institute (ANSI); however, at times during its history it was also known as the American Standards Association (ASA) and the United States of America Standards Institute (USASI). In 1926, under the auspices of the ASA, engineering abbreviations and symbols were standardized. The AIEE, in cooperation with ASA, sponsored in 1928 the development of a glossary of terms used in electrical engineering. This work was coordinated with the IEC[*]. Over the years, ANSI (and its predecessors) has had many responsibilities in the standards arena, including development of standards. Although many people still believe ANSI develops standards, it has not done so for many years. Standards that become American National Standards (ANSs) are written by one of more than 270 standards developers that can submit their completed standards to ANSI for acceptance as ANSs. It is interesting to note that, in the electrical industry, basic standardization was first in order of development, dating back before 1890. Technical standardization came next, with the formation of the Standards Committee of the AIEE in 1898. Manufacturing standardization came only as a result of World War I and did not take effect until 1920.

In Canada, the Canadian Standards Association (CSA) was formed in 1919. It is a not-for-profit organization supported by its members and develops standards in many fields. In 1970, a new organization called the Standards Council of Canada (SCC) was established by an act of parliament to coordinate voluntary standardization in Canada. The SCC is a federal Crown Corporation. The CSA represents Canada on a number of ISO Committees on behalf of the SCC. Standards may be submitted by accredited standards developing organizations to the SCC for approval as a National Standard of Canada.

Regional Standardization. Once standards began to be developed, it did not take long for regional organizations such as the Organization of American States (OAS), the Pan American Standards Commission (COPANT), or the Pacific Area Standards Congress (PASC), and alliances such as the North American Treaty Organization (NATO), to see the value in having common (or harmonized) standards. For example, in May 1923, the OAS (then known as American States of the Pan-American Union) established the Inter-American Electrical Communication Commission (now known as the Inter-American Telecommunication Commission). However, it was not until the formation of three regional standards organizations by the European Economic Community that the world really took notice. In the area of electrotechnology, the European Committee for Electrotechnical

[*]The 6th edition of the IEC Multilingual Dictionary was published in 2005, with over 19,400 definitions in English and French. Equivalent terms wherever available are included in up to 11 additional languages, including Arabic, Chinese, Dutch, German, Italian, Japanese, Polish, Portuguese, Russian, Spanish, and Swedish.

Standardization (CENELEC) was formed in 1973. Telecommunications standardization is the responsibility of the European Telecommunications Standard Institute (ETSI). All other standardization is the responsibility of the European Committee for Standardization (CEN).

Associations. Many associations have come to exist for various reasons, which may include standards. One of the earliest was the Association of Edison Illuminating Companies (AEIC). It was founded in 1885 to provide guidance to the Edison Illuminating companies that were being formed around the United States. The AEIC became a place where problems facing the growing electric utility industry could be solved by pooling the knowledge and experience of managers, engineers, and operators. Most of the work of the AEIC was technical in nature until 1948; however, in 1948 a committee was formed to deal with load forecasting and end-use management. The AEIC today continues to produce standards for equipment, such as cable, used by electric utilities.

Another example of an early association was Aeronautical Radio, Inc. (ARINC), formed in 1929 by four major airlines. ARINC was incorporated to serve as "the single licensee and coordinator of radio communication outside of the [U.S.] government." Once ARINC was organized, the Federal Radio Commission (predecessor of the Federal Communications Commission), transferred responsibility for all aeronautical ground radio stations to ARINC. ARINC continues to provide services today to the airlines, aviation-related companies, and government agencies.

In the field of telecommunications, the Exchange Carriers Association was formed in 1983 as part of the breakup of the Bell System (i.e., AT&T) in the United States. It has recently been renamed the Alliance for Telecommunications Industry Solutions (ATIS). Its membership is open to those involved in telecommunications in North America and the Caribbean. Committee T1 was formed in 1984 to give exchange carriers a voice in the creation of telecommunication standards, which had previously been developed, de facto, by AT&T. ATIS eventually became the secretariat for committee T1; however, committee T1 was retired in 2004 and its standards work was assumed by ATIS.

Standards in Current Times. The information age and global economy have increased the demand for new standards that are internationally acceptable. The completion of the Uruguay Round of negotiations (1986–1994) of the General Agreement on Tariffs and Trade (GATT) led to the establishment of the World Trade Organization (WTO) in 1995 and a new set of agreements covering goods, services, and intellectual property. It also established a new dispute settlement mechanism. The WTO is the only international agency overseeing the rules of international trade with 148 member nations (as of October 2004). In 2001, the WTO began to host a new round of negotiations under the Doha Development Agency. The formation of the European Commission (EC) along with its regional entities and requirements for compliance with European Norms has had a significant impact on standards. For example, in 1990 the European Organization for Certification and Testing (EOTC) was created under a memorandum of understanding between CEN, CENELEC and the European Free Trade Agreement (EFTA) countries. EOTC was formed to promote the mutual recognition of test results, certification procedures, and so forth throughout the EC and EFTA countries. The North American Free Trade Agreement (NAFTA) and EFTA have had similar effects on the standards community.

Programs intended to harmonize standards to make them more internationally acceptable were instituted by many standards developers. Certification (or registration) began to take on additional importance to those organizations that wanted to compete in the global market. As a result, conformity assessment (which includes both registration and certification) programs began to expand, and in a number of instances certification organizations in one country expanded into other countries or partnered with a certification organization in another country. In 1979, ISO established a technical committee to harmonize the increasing international activity in quality management and quality assurance standards. One product of this committee was the ISO 9000 series of standards, which are internationally accepted and can provide a company that uses them with a route to the world markets. After ISO 9000 the ISO 14000 series of standards on environmental management tools and systems were developed. These standards address a company's system for managing its day-to-day operations as they impact the environment.

Other changes that have occurred in the standards development arena are that the standards developers themselves are changing. Many have been renamed to reflect a more international flavor, and most have reengineered their processes to provide standards in a more timely manner or have begun to introduce new products such as emerging technology standards. For example, the IEC introduced the Industry Technical Agreement (ITA) as a new product in 1997 "in its drive to remain relevant in the field of electrotechnology." The ITAs are not produced within the traditional IEC committee structure, nor are they consensus documents; however, they can be produced in months rather than years. Later, the Technology Trend Assessment (TTA) was introduced. A TTA presents the state-of-the-art or trend in a field of emerging technology that might become an area for standardization in the near-to-medium-term. TTAs are typically the result of research or prestandardization work. Although these programs have met with some degree of success, those in the information technology and telecommunications market want standards in place before, not after, their products are created. These industries are more interested in the anticipatory information provided by the standards process than they are in the final standard. They are also interested in new and more flexible forms of standards development.

As a result, these industries have turned to the formation of trade associations and consortia to develop their standards. In the United States, the number of standards produced in this manner will be greater than those produced by traditional SDOs. In contrast to traditional standards which are typically produced by volunteers, trade associations and consortia use paid professionals and provide them with budgets for expenses, research, legal advice, and so forth. Additionally, international standards organizations such as IEC and ISO have recognized a number of these standards as Publicly Available Specifications (PASs). One example of an industry consortium is the World Wide Web Consortium (W3C). It was founded in 1994 and its mission is "to lead the World Wide Web to its full potential by developing protocols and guidelines that ensure long-term growth for the web.

Other examples are: the Unicode Consortium, established in 1991, to bring together leading software corporations and researchers at the leading edge of standardizing international character encoding; the Open Group, established to answer questions in IT that corporate IT users need answers to by aiding in the development and implementation of a secure and reliable IT infrastructure; the DSDM Consortium, established in 1994 to develop and promote a public-domain rapid application development method; and the Northeast American Electric Reliability Council (NERC), founded in 1968 after the Northeast Blackout to promote reliability of the electrical supply for North America.

28.3 STANDARDS AND THE LAW

Voluntary Standards System. The word *standard* has a number of meanings, but in the context of trade or engineering, it refers to voluntary technical standards that are normally developed by a consensus of experts. Many standards set safety or performance requirements for products or services, for example, standards for workshoes for those involved in electrical work to be "nonconductive with a reinforced toe" for safety purposes. These standards are not normally developed by lawmakers, and because they are outside of the mandatory scope of laws, they are sometimes referred to as *voluntary standards*. In general, standards are developed under a voluntary system. To the extent that their adoption is also voluntary, there is less vulnerability to legal liability. However, many standards are made mandatory, either through reference in purchase specifications and contracts or through adoption by government bodies as regulatory documents. For example, in certain states it may be illegal for a person to drive (or ride as a passenger) a motorcycle without a helmet that meets the requirements of a specific standard. Under such circumstances compliance ceases to be voluntary and the effect of the document is to disqualify or limit the acceptability of certain products or services.

The ability of standards to limit acceptable suppliers is a potential danger of standards and the processes under which they are developed must minimize the possibility of discrimination against specific companies. One can begin to see that standards could be developed that contain absurd requirements that could act as a barrier to trade from a foreign nation, or within the same nation violate antitrust laws. Additionally, once a standard is written into the law, if the law simply states that compliance with standard *xyz* is required, then any revision to the standard (as occurs periodically) has the effect of amending the law.

Legality of Standards. The legality of standards activities is primarily affected by laws related to the fixing of prices, conspiracy in restraint of trade, and intellectual property. Throughout history, standards have been well-known barriers to trade as countries hide protectionism in the veil of an absurd standard. An example would be a standard written by a country that requires the use of a specific material available only in that country for a particular part of the product. Today, however, standards are covered by the GATT, the WTO, and other trade agreements such as NAFTA. The WTO supports the use of international standards developed under the auspices of international standards organization such as IEC and ISO. One reason for this is the belief that it is felt that the international development process will identify and exclude any documents that contain hidden trade barriers. From the users' perspective, certification to international standards should result in greater international acceptance for their product or service.

In the United States, the two key governmental agencies involved are the Federal Trade Commission (FTC) and the Department of Justice. In 1975, the FTC and the Justice Department held hearings on a number of abuses of standards development and certification activities. These abuses involved individuals involved in the standards process who attempted to use standards for market advantage or to deny competitors entrance into an established market. One conclusion of the hearings was that there needed to be some fundamental guidelines and practices that would ensure that the activities related to standards development would be "fair." Those fundamentals were due process, openness, balance, public notice, and the right to participate and appeal. These are the same basic principles that govern standards development around the world; however, they have, in many instances, been interpreted and reinterpreted to the point where they slow down the process. This, as stated earlier, has driven many away from the traditional standards organizations, especially in the areas of new technology.

Since standards activities involve meetings in which representatives of competing organizations make agreements that affect engineering and industrial practices (both of which have economic implications), such meetings must take place under conditions which are subject to carefully regulated procedures. Failing this, participants could be subject to charges of violation of antitrust or conspiracy statutes. Trade associations and consortia are particularly vulnerable in this respect, as meetings restricted to their membership involve participants who tend to be exclusively competitive manufacturers, whereas meetings of committees of professional societies involve technical personnel who are more apt to be representative of the total industry (both manufacturers and users), independent consultants, government personnel, educators, and scientists. Similarly, international standards committees are populated by experts from the national committees from the member countries. These persons individually represent the consensus of experts in their country when developing or voting on standards. However, the degree of liability of participants in standards development activities is virtually negligible when these activities are conducted under the auspices of, and under the strict rules of, an organization experienced in standards development, that is, an organization whose procedures are designed to promote fair and unprejudiced participation by all eligible parties.

Certification. The certification of a product provides additional assurance that a product is reasonably safe and reasonably suited for its intended function. Certification is particularly important for products that are purchased by the general public (i.e., consumers). Legal action against certifying organizations is rare; however, negligence in the certification process could cause a certifying organization to become exposed to a claim for liability.

Patents. The issue of patents as related to their use (or specification) in standards has become something that most standards developers have had to deal with, particularly for new technology standards. Most standards developers have patent policies that require disclosure of patents at the time a proposal for development of a standard is submitted. They further require disclosure at any time in the standards development process that it becomes known that a patent is applicable to a standard. Additionally, the patent holder is normally required to provide a letter stating that (1) the patentee will not enforce any of its present or future patent(s) whose use would be required to implement the proposed standard against any person or entity using the patent(s) to comply with the standard or (2) a license will be made available to all applicants without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination. If these conditions are met, the patent may be included in the standard; however, the standards developer normally publishes a disclaimer in the standard making no claims as to the validity of the patent or the reasonableness of rates and/or terms and conditions of the license.

Personal Liability. An area of legal concern for participants in the standards generation or approval process is the question of legal liability. A typical situation deals with the case where an accident occurs under circumstances where potentially negligent parties demonstrate that they faithfully complied with the provisions of the applicable safety standards. The question here is one of the extent of liability of those who participated in the generation or adoption of the standard. A somewhat equivalent situation arises in product liability cases. Any such claim in a legal action turns on allegations of negligence in writing the standard. The general conclusion held by counsel is that members of voluntary standards committees operating under procedures that embody the fundamental principles of due process, openness, balance, public notice, and the right to participate and appeal are not likely to incur significant legal risks.

Some standards developers indemnify those persons who are members of the organization, provided the processes of the organization have been followed during the standards development process. Other standards developers require those participating on its standards writing groups to sign a statement attesting to the fact that they will follow the organization's procedures when participating in standards activities for the organization.

28.4 THE VOLUNTARY STANDARDS PROCESS

Voluntary Development of Standards. There are literally tens of thousands of experts in the fields of electrotechnology, telecommunications, and information technology who participate in standards development worldwide. It should be clear that it is the process by which the standards are developed that is of importance. The process should embody the fundamental principles; however, there are almost as many processes as there are standards developers. Although it would be impossible to describe them all in this section, the information age has made access for interested parties easier than ever before. Today, the World Wide Web allows direct access to a wealth of standards information. A listing of organizations (including their acronyms) and their websites, is provided at the end of this section to enable an interested party begin exploration of the "world of standards." Many of the websites contain the full text of the procedures followed by the standards developer. These can frequently be downloaded.

Although the development processes are designed to enable the broadest possible input to standards, duplicate standards, conflicts between standards, and other problems may exist. The resolution of these problems may be by the standards developers themselves, or by a national or international organization. For example, in the field of electric power cable, conflicts have arisen between standards developed by ASTM and IEEE that had to be resolved by the developers, or by ANSI where resolution was not achieved by the developers. In the case of duplicate standards in a voluntary system, the standard of choice will usually be decided by the market (i.e., the standard referenced by users will become the standard used). Complementary standards for a product are sometimes developed by two or more standards organizations, with each organization responsible for preparing standards within its area of expertise. For example, for power switchgear standards in the United States, IEEE develops those standards related to specifications and tests, while NEMA is responsible for those standards related to ratings. Occasionally, two or more organizations may develop a standard jointly; however, issues related to copyright and reproduction need to be agreed to by each organization before work begins. An example of this is standard IEEE/ASTM S10 which was jointly developed by IEEE and ASTM International and first published in 1976. This standard is intended to give authoritative information on SI and guideance on its application to U.S. citizens and industry. Its use became more important in 1988 when the U.S. Metric Conversion Act was amended to designate "the metric system of measurement as the preferred system of weights and measures for U.S. trade and commerce."

Approval of standards is normally by consensus, with the definition of what constitutes consensus defined by the developing organization's procedures. What this means is that a standard may be made available to users once consensus is reached by the sponsor. A problem here is that some sponsors apply a more rigid interpretation of consensus than intended (some even try to attain unanimity), which ultimately delays the standard. In these instances, the sponsor is doing a disservice to the users that need the standard. Although the sponsor is responsible for the technical content of the standard, the standards developer often has some sort of authority (e.g., a standards board) that is responsible for ensuring the organization's process has been followed. Approval by that authority is the final step in the process before publication. Here, too, interpretation of the process by the authority can cause standards to be returned to the sponsor, resulting in an unnecessary delay of the standard.

Another part of the process that has not been discussed is the maintenance of a standard once it is developed. All standards require periodic review to ensure that the information contained in them is current. Once reviewed, the standard may be reaffirmed (or confirmed), revised, or withdrawn. The standards developers also have the ability to administratively withdraw a standard if it has not been reviewed by its sponsor within some maximum period of time. For example, in the United States, the maximum time a standard can exist without being revised or reaffirmed is 10 years. The IEEE requires its standards to be reviewed every 5 years. The sponsor may request an extension of up to 5 years with a explanation (a 2-year extension is normally granted on request, without explanation). The standard is administratively withdrawn by IEEE after 5 years if no extension is requested and after 10 years with extensions. Internationally, both ISO and IEC require a review of their standards by the responsible Technical Committee (or Subcommittee) at least once every five years. The process for maintenance of standards developed by some consortia may be less than adequate and may be a future issue in standards. This is an area that traditional standards developers are targeting for new work (i.e., to provide standards maintenance services to consortia).

28.5 TERMINOLOGY IN STANDARDS

Standards Terms. The following comprises a partial list of terms used by participants in standards activities. Many of these terms have unique and specialized meaning when used in the context of standardization, and a brief definition is given for each as applied in this context. Some of the terms are specific to the processes of the IEC and are identified by an (IEC) at the end of the explanation of the term.

approval stage The point after the enquiry stage, at which the final-draft international standard (FDIS) is circulated to the National Committees for a 2-month voting period. If the FDIS is approved it is published, and if it is not approved it is referred back to committee for reconsideration (IEC).

balance The characteristic of a standards approving unit (committee, subcommittee, or working group) which assures that all classifications of interests are represented and that no single classification has a representation sufficiently large to enable it to unduly influence the resulting output.

balanced committee A committee so constituted as to maintain a balance among its members. Many committees are balanced among manufacturers, users, and general-interest classifications.

basic standard A standard common to all disciplines, or to an overall technology.

canvass A method used for approval of standards which is dependent on circulation of a draft document to a list of concerned organizations for review and ballot.

*American National Standard for use of the International System of Units (SI): The Modern Metric System.

certification An attestation to the effect that a particular product or service meets the requirements of a relevant standard.

certification mark A special kind of trademark that appears *only* on products that have been certified against a standard.

classification of membership The classification assigned to a participant or member of a standards developing unit which identifies the member's functional relationship or interest in the subject to be standardized. Thus a participant may be a manufacturer of a product being standardized, a user or purchaser of the product, a technically qualified expert with no well-defined functional relationship (classified as general-interest), a labor or insurance representative (in the case of safety standards), or a constructor (one who installs the product for use by others). A variety of other classifications is possible as dictated by the scope of the standards activity.

code (*a*) A body of recommendations of good practice to be followed during design, manufacture, construction, installation, operation, and maintenance to satisfy considerations of safety, quality, economy, or performance in a given application. (*b*) A particular form of identification marking or reference which serves the dual purpose of establishing in a systematic manner the complete identity of an individual product and of identifying its similarity with other products. It may consist of a brief, systematic combination of letters, numerals, and symbols.

committee stage The point at which the document is submitted to the National Committees for comment and ballot if the document is intended to be published as a technical report (IEC).

conformity assessment An activity or set of activities that determines directly or indirectly whether relevant local product requirements have been fulfilled. Typical forms of conformity assessment include testing, inspection, assessment, auditing, certification, registration, and certification.

consensus A substantial agreement of those concerned. It implies that no important interested parties are strongly opposed on substantive grounds, or alternatively, that any opposition is in a small minority and the changes required to effect agreement by this minority would lead to substantive disagreement by the majority. Consensus implies that all disagreements have been given careful consideration and all reasonable attempts have been made for their resolution.

designation A definite and distinguishing name or symbol given to a product or to a group of functionally similar products or to an abstract matter. It emphasizes the group similarity but does not bring out the differences among the various members of the group.

dimensional interchangeability A condition in which the dimensions of two or more products are such that one can physically replace another in a given application.

dimensional standard A standard whose main content is dimensions and sizes of a product or group of products.

e.ballot A "letter ballot" or equivalent that is conducted electronically (e.g., via e-mail).

enquiry stage A point before the approval stage where the bilingual Committee Draft for Vote (CDV) is submitted to P-members of a technical committee for a 5-month voting period. It is the last stage at which technical comments can be taken into consideration (IEC).

functional interchangeability A condition where the characteristics of two or more products are such that they are able to perform the same functions.

guide A standards document that provides alternative information which comprises good engineering practice. Guides may contain application information for use of products and may be tutorial in nature. The user should be cautioned that the use of the word "guide" in the title of a document does not guarantee that the document is in fact nonmandatory. There are many governmental regulatory guides which in fact set forth mandatory requirements. Conversely, many documents that are differently titled are in fact guides.

harmonization The act of coordinating requirements from multiple standards (e.g., multiple countries or multiple SDO's standards) and copublishing the resulting document.

harmonization committee (or task force) A group of individuals responsible for technically developing the proposed draft of the harmonized standard. The group typically consists of a representative of each involved country (or SDO), the secretariat, the chair.

interface standard A standard whose main purpose is to ensure coordination between systems.

international standard A standard that has been adopted by a recognized international standards body (such as IEC or ISO).

joint publication A standard that has been submitted through the standards development process of two or more SDOs and is published separately by all involved SDOs. A joint publication may also be referred to as a harmonized or copublished standard, or if the SDOs represent different countries, as a binational or trinational standard.

letter ballot A ballot used in standards development to determine agreement on a draft standard, or to generate comments that will be instrumental in developing a document on which consensus agreement can be achieved. Such ballots provide for affirmative and negative votes. Negative votes, however, must be accompanied by reasons in sufficient detail to enable the writers of the document to determine what steps need be taken in revision to change the vote from negative to affirmative. The primary advantage of a letter ballot is that it provides adequate time for the recipients to review thoroughly the document which is subject to ballot.

marking The action and the result of stamping, inscribing, printing, or labeling marks, symbols, letters, or numerals on a product or its package for the purposes of identifying the product.

may An operative verb used in a standards document which identifies a possible means for satisfying a requirement. For example, several alternative procedures may be indicated for measuring a particular characteristic or phenomenon, and the selection of the most suitable procedure is left to the user of the document.

national standard A standard that has been adopted by a recognized national standards body (such as ANSI), or a standard that is in effect recognized and used nationally in preference to other documents.

O-member An observer member of a technical committee who has the right, but not the obligation, to vote and attend meetings (IEC).

performance characteristic A characteristic of a product which determines the product's suitability for a specific application.

P-member A participating member of a technical committee who is obliged to attend meetings and vote on a draft international standard (DIS, IEC).

preliminary stage Projects envisaged for the future but not yet ripe for immediate development, or preliminary work, such as better definition of a project for new work, data collection, or round-robin tests necessary to develop standards, which is not part of the standardization process (IEC).

preparatory stage The phase during which a working draft (WD) of a document is prepared (IEC).

product standard A standard containing requirements to be met by a product or group of products, usually including, directly or by reference to other standards, all or some of the following elements: dimensions, performance characteristics, other characteristics, and test methods.

proposal stage A proposal for new work originated from industry via a National Committee, communicated to the members of a technical committee or subcommittee with a form. A simple majority vote takes place within 3 months; if the result is positive and five members agree to actively participate, it is included in the work program (IEC).

rating A characteristic of a product which is determined in an arbitrary, yet consistent, manner, based on the intended function of the product.

recommended practice A standards document that provides information on good engineering practice. Such documents may contain application information for use of products.

safety standard A standard whose primary purpose is to ensure the safety of people and property.

SDO Standards Developing Organization.

secretariat An organization that assumes the responsibility for providing administrative oversight of a standards committee's activities and assures compliance with all applicable procedures.

self-certification An attestation by a manufacturer or supplier of a product or service that it meets the requirements of a relevant standard.

shall An operative verb used in a standards document which indicates a mandatory requirement that must be specifically complied with for conformance to the document.

should An operative verb used in a standards document which indicates a problem area that must be resolved and specifies a requirement, compliance with which resolves the problem. In this sense, the verb "should" can be read as "shall." Alternatively, it is allowable under the document to use some other method which can be proved to resolve adequately the condition or problem area addressed. In some cases, it is also possible to demonstrate clearly that the condition or problem area addressed does not in fact exist, or apply to the product or circumstance in a specific instance.

simplification A form of standardization consisting of the reduction of the number of types of products within a definite range to that number which is adequate to meet prevailing needs at a given time.

specification A standards document that specifies all the characteristics and conditions to be met by a product or service to be supplied to the purchaser. Such a document may refer to other standards, selecting among the specific allowable options. A specification is intended to be a complete purchasing document.

sponsor The group (e.g., a technical committee) that assumes responsibility for the development and/or maintenance of a standard.

standard A documented agreement containing technical specifications or other precise criteria to be used consistently as rules, guidelines, or definitions of characteristics to ensure that materials, products, processes, and services are fit for their purpose.

standardization An activity aimed at an increase of order, giving solutions for recurring problems in the spheres of scientific, technological, and economic activity. Generally it consists of the processes of formulating, issuing, and implementing standards.

terminology standard A standard containing exclusively terms and their definitions.

test standard A standard containing test methods which may be combined with other requirements related to testing, such as sampling, use of statistical methods, and sequence of tests.

third-party certification An attestation by a recognized, technically qualified, independent organization that a product or service supplied by others meets the requirements of a relevant standard. Such certification may be based on inspections and tests conducted by the certifying organization, or on supervision, monitoring, or auditing by the organization of such tests which may be conducted by others. The tests may be performed by the manufacturer or supplier of the service or product while being witnessed or audited by the certifier.

trial-use A publication (standard, recommended practice, or guide) that is effective for not more than 2 years, during which time comments and criticisms from a broad constituency are sought. In the absence of comments during the trial-use period, the document is subject to automatic approval on recommendation of the sponsor.

unification A form of standardization in which two or more specifications are combined into one in such manner that the products obtained are interchangeable in use.

28.6 ISO 9000 AND ISO 14000 STANDARDS

ISO 9000 and ISO 14000 Overview. Unlike the vast majority of ISO standards which are highly specific, ISO 9000 and ISO 14000 are generic families of standards and guidelines relating to management systems. When introduced in 1987 and 1997, ISO 9000 and ISO 14000, respectively, introduced standards to a much broader business community (i.e., beyond organizations in the field of engineering and science only) than did any standards published previously. ISO 9000 is concerned primarily with quality management systems, while ISO 14000 is concerned mostly with environmental management systems. In the context of the standards, *quality management* means what an organization does to conform to its customers' requirements; *environmental management* means what an organization does to minimize harmful effects on the environment caused by its activities. Neither standard is a product standard, and organizations such as law firms, consulting engineers, or standards developers can become ISO 9000–certified. The reason for this is that both ISO 9000 and ISO 14000 are concerned with the way an organization goes about its work (i.e., the process), rather than the direct result of the work (e.g., a product or a service). (For additional information visit the ANSI, ASQ, ISO, and NIST Websites listed at the end of this section.)

ISO 9000 History. In 1959, the U.S. Department of Defense (DoD) established a quality management program designated as MIL-Q-9858, which was later revised to MIL-Q-9858A. NATO essentially adopted the provisions of MIL-Q-9858A in 1968, and published them in Allied Quality Assurance Publication 1 (AQAP-1). In 1970, the U.K. Ministry of Defense adopted the provisions of AQAP-1 when it published its Management Programs Defence Standard DEF/STAN 05-8. In 1979, the British Standards Institution (BSI) developed the first commercial quality management system standard, BS 5750. It was from all of these documents, and BS 5750, in particular, that ISO created the ISO 9000 standards family of documents (the base standards of which are shown below).

When Countries Adopt ISO 9000. By 1992, the European Economic Community (EEC) and 56 countries had adopted ISO 9000. The EEC and other countries assigned numbers to the adopted ISO 9000 standards according to their own national standards numbering system. In the United States, the ISO 9000 series was adopted as ANSI/ASQ Q9000 in 1987. The ANSI/ASQ Q9000 series is essentially identical to the ISO 9000 series with the exception that the text incorporates customary American English language and spelling. Some other examples of adoption of the ISO 9000 series are the EEC as the European Norm (EN) 29000 series, the United Kingdom as BS 5750 Parts 0 to 3, Pakistan as PS 3000-3004 series, Tanzania as TZS 500-504, and China as GB/T 10300.1–10300.5.

Certification. ISO does not itself certify conformity to ISO 9000. This is done by independent certification bodies in different countries. There is also no "official" database of entities certified to ISO 9000.

ISO 9000. This standard, entitled *Quality Management Systems—Fundamentals and Vocabulary*, explains fundamental quality concepts. Additionally, it defines terms and provides guidance on selecting, using, and tailoring the standards in the series.

ISO 9001. This standard, entitled *Quality Management Systems—Requirements*, is the most comprehensive standard in the series. It addresses all the elements in design, development, and so on and provides requirements for quality planning.

ISO 9002. Superseded by the 2000 Edition of ISO 9001.

ISO 9003. Superseded by the 2000 Edition of ISO 9001.

ISO 9004. This standard, entitled *Quality Management Systems—Guidelines for Performance Improvements*, provides guidance for developing and implementing an internal quality system.

Since their introduction, the number of standards in the ISO 9000 and ISO 14000 families have grown to more than 40 documents. In 2004, ISO 90003 was introduced which addresses guidelines for the application of ISO 9001 to computer software.

28.7 INTERNATIONAL ORGANIZATIONS

The International Electrotechnical Commission (IEC). The IEC, a nongovernmental body located in Geneva, Switzerland, is the world organization that develops and publishes international standards for electrotechnology and related technology. Its membership is limited to countries. Today, the IEC membership consists of almost 66 participating countries, including all the world's major trading countries (see Table 28-1). Full membership in the IEC allows the country to participate in all international standardization activities. Participation of a country is by a national committee. Each national committee agrees to open access and balanced representation from all electrotechnical interests in its country (i.e., public and private). A country may also become an associate member, which allows for limited participation in the IEC. Associate members have observer status at all IEC meetings, but have no voting rights.

The mission of the IEC is to promote, through its members, international cooperation on all questions of electrotechnical standardization and related matters, such as the assessment of conformity to standards, in the fields of electricity, electronics, and related technologies, including magnetics and electromagnetics, electroacoustics, telecommunication, energy production and distribution, terminology and symbols, measurement and performance, dependability, design and development, safety, and the environment.

The work of the IEC is carried out by more than 10,000 experts worldwide who participate on more than 200 technical committees and subcommittees, and more than 700 working groups. All IEC publications are bilingual (English and French). Certain documents have also been translated into Spanish. The Russian Federation National Committee develops Russian-language editions of IEC documents. English, French and Russian are the three official languages of the IEC; however, in all bodies of the IEC other than the Council, discussions may be held in English and/or French, following agreement of the delegates. Standards developed by the IEC follow the procedures in the ISO/IEC Directives, published jointly by both organizations and administered by the Joint Technical Program Committee (JTPC). These directives are published in two parts. Part 1 covers the procedures

TABLE 28-1 Member Countries of the IEC

Argentina	Iceland*	Pakistan
Australia	India	Republic of the Philippines†
Austria	Indonesia	Poland
Belarus	Iran	Portugal
Belgium	Ireland	Romania
Bosnia & Herzegovina*	Israel	Russian Federation
Brazil	Italy	Saudi Arabia
Bulgaria	Japan	Serbia and Montenegro
Canada	Kazakhstan*	Singapore
China	Kenya	Slovakia
Colombia*	D.P.R. of Korea*	Slovenia
Croatia	Republic of Korea	South Africa
Cyprus*	Latvia*	Spain
Czech Republic	Lithuania*	Sweden
Denmark	Luxembourg	Switzerland
Egypt	Macedonia* (former Yugoslav Rep. of)	Thailand
Estonia*	Malaysia	Tunisia*
Finland	Malta*	Turkey
France	Mexico	Ukraine
Germany	Netherlands	United Kingdom
Greece	New Zealand	United States of America
Hungary	Norway	Vietnam*

*Associate member.
†Suspended 1-13-2003.

for the technical work. Part 2 covers drafting and presentation of international standards. These common procedures were adopted by ISO/IEC in recognition of the need to develop timely and cost-effective international standards. The joint ISO/IEC Committee JCT 1, Information Technology develops and maintains its own procedures. All of these procedures may be accessed on each organization's web site.

Figure 28-1 is an organization chart for the IEC. The Council, which is a "general assembly" of committees, is the supreme authority of the IEC and sets policy, financial objectives, and strategy. The Council delegates the management of all IEC work to the Council Board. The responsibility for standards and conformity assessment is assumed by the Standardization Management Board and Conformity Assessment Board, respectively. In the areas of standards and conformity assessment, the IEC works closely with other international organizations such as the ISO, WTO, and ITU. It also has relationships with governmental agencies and regional standardization organizations. For example, an agreement between IEC and CENELEC ratified in 1996, known as the Dresden Agreement, addresses common planning of new work and parallel IEC/CENELEC voting. Other standards organizations may interface with the IEC as liaisons. For example, the IEEE is recognized as a Class D liaison to the IEC.

International Organization for Standardization (ISO). Like the IEC, ISO is a nongovernmental body, located in Geneva, Switzerland, that is a worldwide federation of national standard bodies. ISO is the world's largest developer of standards, publishing approximately 12,000 documents. Its membership is limited to countries, and at present, some 146 countries are members of ISO. The member body of a country to ISO is the national body "most representative of standardization in its country," and only one body in each country can be admitted to membership in ISO. So, for example, ANSI is the ISO member from the United States, Standards Council of Canada is the ISO member from Canada, Standards Australia is the ISO member from Australia, and DIN is the ISO member from Germany. ISO has two other categories of membership for countries. The first is correspondent member, for countries which do not yet have a fully developed national standards activity; the second is subscriber member, for countries with very small economies.

ISO likes to think of its documents as international agreements that are published as international standards. Some might note an apparent inconsistency between the short form of the organization's name, specifically, ISO and its official title. That is because the short form of the name is not an acronym as many believe, but rather a word ISO derived from the Greek *isos* meaning equal. On this basis, the connection between "standard" and "equal" is easy to understand.

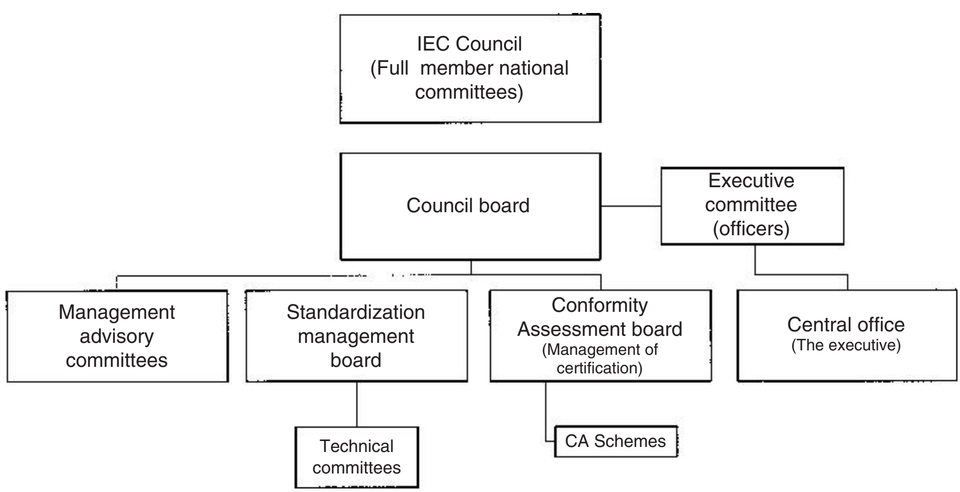


FIGURE 28-1 IEC structure and management.

The mission of ISO is "to promote the development and standards related activities in the world with a view to facilitating the international exchange of goods and services, and to developing cooperation in the spheres of intellectual, scientific, technology and economic activity." The scope of ISO's work is not limited to any particular field, except that electrotechnology is the responsibility of the IEC and information technology is carried out by JTC 1 (a joint committee of ISO/IEC).

Among the documents that affect those in electrotechnology that ISO is responsible for is the universal system of measurements, known as SI (Système International d'Unités) units, which are described in a series of 14 international standards. The work of ISO is carried out by more than 30,000 experts worldwide who participate on more than 700 technical committees and subcommittees, and more than 2000 working groups. An organization chart for ISO is shown in Fig. 28-2. ISO cooperates with other international bodies such as the IEC and ITU, with regional standardization organizations and is building a strategic partnership with the WTO. ISO also has liaisons with almost 550 entities worldwide that are interested in specific aspects of its standardization work. All ISO publications are bilingual (English and French) and are developed following the ISO/IEC directives.

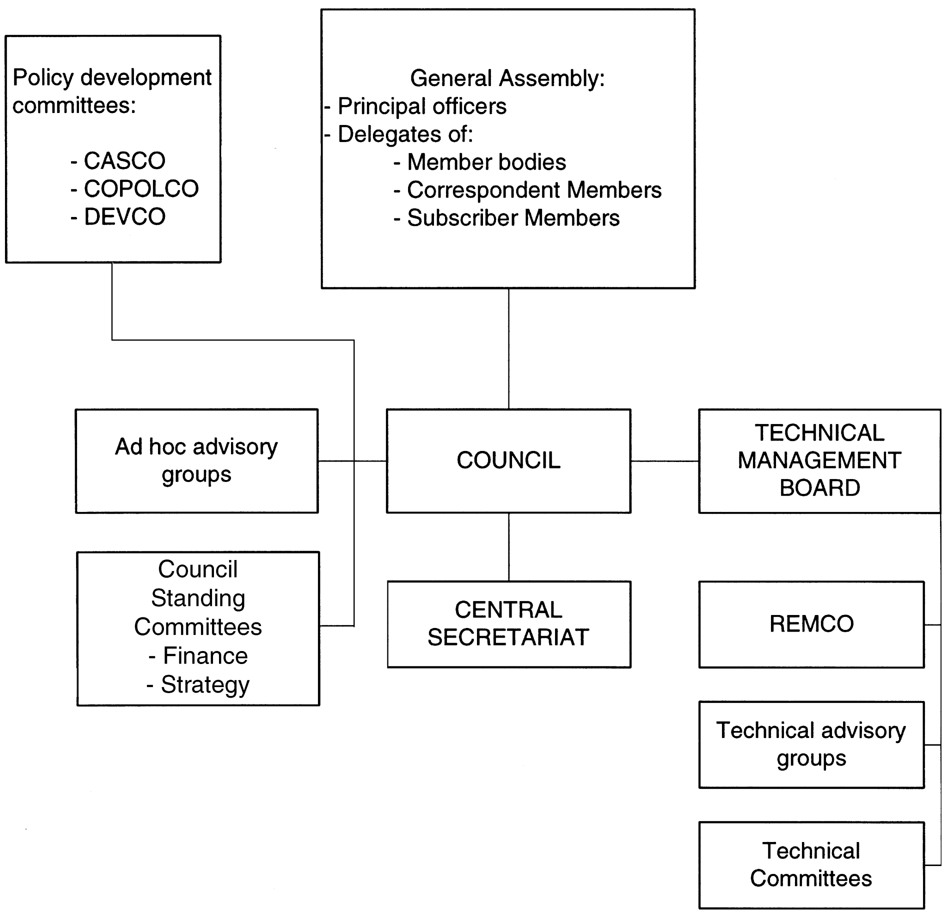


FIGURE 28-2 The ISO organization. (*Legend:* CASCO—Committee on Conformity Assessment; COPOLCO—Committee on Consumer Policy; DEVCO—Committee on Developing Country Matters; REMCO—Committee on Reference Materials.)

The International Telecommunication Union (ITU). The ITU, unlike the IEC and ISO, is part of the United Nations (UN), and its members are governments. It is located in Geneva, Switzerland and has 189 member states (i.e., countries). The ITU is an organization "within which governments and the private sector coordinate global telecom networks and services." It is the leading publisher of telecommunication technology, regulatory, and standards information. The governing bodies of the ITU are shown in Fig. 28-3. The Plenipotentiary Conference has the responsibility to deterrmine the structural and operating changes necessary in the ITU to effectively serve the requirements of the international telecommunications community.

Joint Technical Committee 1 on Information Technology (JCT 1). In recognition of the broad scope and applicability of international standards in the field of information technology, ISO and the IEC approved an agreement in 1987 that formed a joint committee, known as JTC 1 to carry out this work. Membership in JTC 1, like ISO and IEC, is limited to countries; its publications are bilingual (English and French). JTC 1 has 27 member countries (P-members) and 39 observer countries (O-members). It has 17 subcommittees that accomplish its work, and even though it is part of the ISO/IEC, it has its own Website.

The International Committee on Illumination (CIE). The CIE, located in Wien (Vienna), Austria, is "an organization devoted to international cooperation and exchange of information among its member countries on all matters relating to the art and science of lighting." Its short form, CIE, is an acronym from its French title, Commission Internationale de l'Eclairage. The subjects covered by CIE in the fields of light and lighting include vision, photometry, colorimetry, application of light both indoors and out, environmental effects, aesthetic effects, and means for the production and control of light and radiation. The spectrum of light covered is both natural and synthetic over the ultraviolet, visible, and infrared regions. CIE also addresses the optical, visual, and metrological aspects of the processing and reproduction of images using all types of analogous and digital imaging devices, storage media and imaging media. Participation in CIE is like that of other international organizations, by national committee, with 38 countries represented at present. In addition to its own publications, the CIE has published standards jointly with IEC (e.g., the IEC/CIE International Lighting Vocabulary) and ISO.

The Internet Society (ISOC). The ISOC, located in Reston, Va. (U.S.A.), is a nonprofit, nongovernmental, international professional membership organization. It was formed to provide an institution home for the Internet Standards process, as well as to provide financial support for the process.

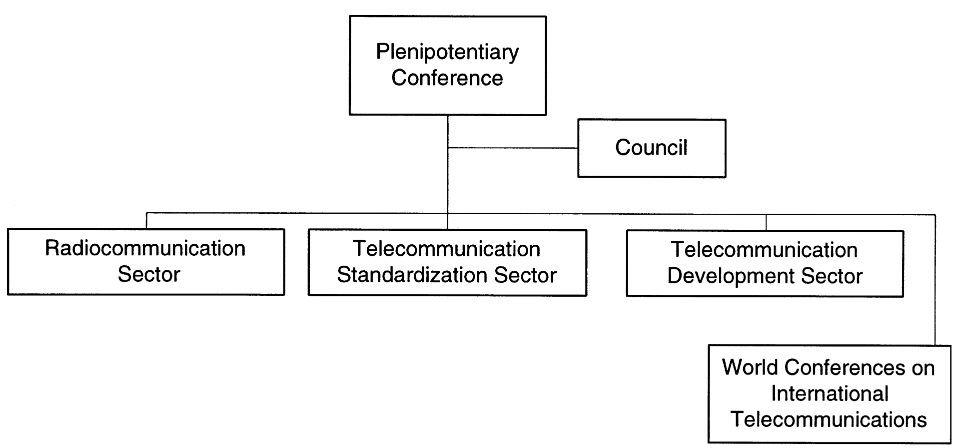


FIGURE 28-3 Governing bodies of the ITU.

It provides the leadership for addressing issues that affect the future of the Internet and it is the organizational home for the groups responsible for Internet infrastructure standards. The combined efforts of the ISOC and IETF have formed and maintain the foundation for all modern networks and Internet products and services. The primary organization supported is the Internet Engineering Task Force (IETF), which since its inception had been supported primarily from research supporting agencies of the U.S. government. The ISOC also charters the Internet Assigned Numbers Authority (IANA) as the central cleaning house to assign and coordinate unique parameter values for Internet protocols. Today, the ISOC has more than 100 organizational and 20,000 individual members in more than 180 countries.

The World Trade Organization (WTO). The WTO, located in Geneva, Switzerland, is an intergovernmental organization and is the only international organization involved in the rules of trade between nations. The WTO is a place where member governments can go to try and resolve trade problems that are between them. The main decision-making bodies are councils and committees consisting of the WTO's entire membership. The authority of the WTO comes from agreements, which are the legal ground rules for international trade and commerce policy. These agreements have three main objectives: "to help trade flow as freely as possible; to achieve further liberalization gradually through negotiation; and to set up an impartial means for settling disputes." Prior to the establishment of the WTO, the GATT had governed international trade and commerce since 1948. The GATT itself was initially an agreement that was provisional. It later became an international organization created to support the agreement but was not recognized in law as an international organization; however, the GATT (i.e., the agreement) has been incorporated into the WTO agreements. Unlike the GATT, which only dealt with trade in goods, the WTO agreements deal with other issues such as copyright, trademarks, patents, industrial designs, and trade secrets. Members of the WTO agree to use international standards (or harmonized standards) to minimize the risk of barriers to trade being introduced by standards. The WTO rules, however, sometimes support maintaining trade barriers when necessary such as in cases where there is a need to protect consumers or to prevent the spread of disease.

The World Intellectual Property Organization (WIPO). The WIPO, located in Geneva, Switzerland, is one of the United Nation's specialized agencies. As such, it is an intergovernmental agency, dedicated to "promoting the use and protection of works of the human spirit." It is responsible for "the promotion and protection of intellectual property throughout the world through cooperation among States, and for the administration of various multilateral treaties dealing with the legal and administrative aspects of intellectual property." The WIPO currently has 182 member states and administers 23 international treaties. Although the WIPO officially became part of the UN in 1974, its "roots" go back to 1883 when the Paris Convention for the Protection of Industrial Property was the first treaty designed to protect intellectual creations through patents and trademarks. Protection of copyrights was added in 1886 with the Bern Convention for the Protection of Literacy and Artistic Works. These groups united in 1893 to form what became the WIPO in 1970.

28.8 REGIONAL ORGANIZATIONS*

European Committee for Electrotechnical Standardization (CENELEC). CENELEC, located in Brussels, Belgium, is a not-for-profit organization and has been officially recognized as the European Standards Organization in its field by the EC in Directive 83/189 EEC. Its membership is made up of national committees from 28 European countries. More than 40,000 technical experts are

Author's note: It is not possible, nor practical, to list all the regional and national organizations, and other standards organizations, that are involved with electrotechnology, telecommunications, and information technology standards. For those readers having additional interest, the World Wide Web provides one with an excellent means of researching these organizations. The reader is also referred to the listing of Websites at the end of this section.

involved in the development of its standards, which are called European Norms (ENs). ENs in the field of information technology are developed and published jointly with the European Commission for Standardization (CEN).

In developing standards, it is the policy of CENELEC to use an IEC standard if it exists (approximately 80% of CENELEC ENs have been developed in this manner). When new work is proposed, it is offered to the IEC with a request that it be undertaken at the international level. CENELEC will develop a standard only if the IEC does not want to undertake the work or if the IEC cannot meet the target dates for the standard. If the IEC assumes the development of a standard, parallel IEC/CENELEC procedures and voting (see Fig. 28-4) designed to result in identical international standards and ENs are used. Once CENELEC has begun work on a standard, or selected an IEC document, all national work (i.e., for CENELEC members) on the same subject is immediately stopped. This suspension of national work for CENELEC members is called "standstill."

Voting in CENELEC is weighted according to the size of the country. Larger countries such as Germany, Italy, France, and the United Kingdom have 10 votes. The smaller countries have one or two votes depending on size. An affirmative vote must meet two conditions. First, a majority of the national committees must vote affirmative and second, at least 71% of the weighted votes must be affirmative. Once an EN is approved it is mandatory that it be adopted as a national standard by CENELEC members (this is unlike most standards processes, e.g., those of the IEC, which allow for

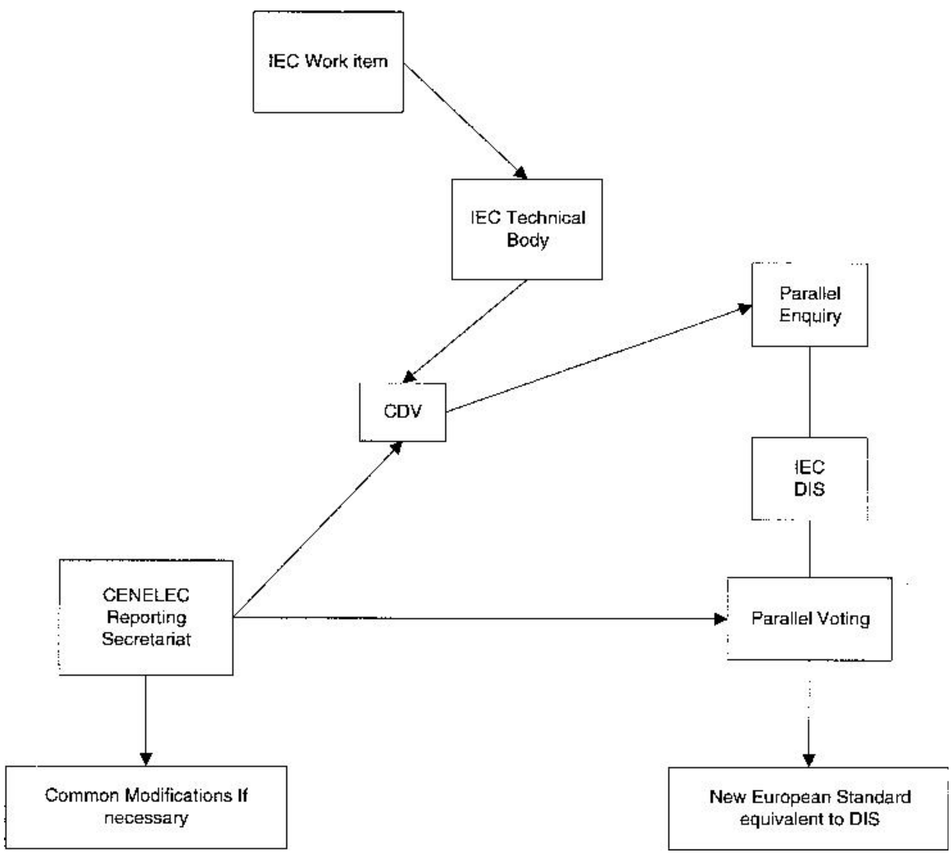


FIGURE 28-4 IEC/CENELEC parallel procedures for standards development. (*Legend:* CDV—Committee Draft with Vote; DIS—Draft International Standard; IEC—International Electrotechnical Commission.)

voluntary adoption). Each CENELEC member is also obligated to withdraw any of its national standards that conflict with an approved EN. On this basis, it is acceptable for a national committee to vote one way in CENELEC (e.g., negative) and another way in IEC (e.g., affirmative).

European Telecommunications Standards Institute (ETSI). ETSI, located in Sophia Antipolis, France, is a nonprofit organization whose mission is "to produce telecommunications standards for use today and the future." It is an open forum (i.e., to European organizations) that brings together some 688 members from 55 countries, representing governments, network operators, manufacturers, service providers, and users. ETSI, unlike CEN and CENELEC, produces voluntary standards, some of which may be adopted by the EC as the technical base for directives or regulations. Promotion of international standardization is favored by ETSI, and as such, it coordinates its activities with international bodies (primarily with the ITU). At present, more than 3500 experts are working in more than 200 groups in ETSI, with more than 13,200 documents already published.

Voting on standards within ETSI is done according to its Rules of Procedure depending on the issue, with only those that are full and associate members allowed to vote. The voting method is weighted individual voting.

28.9 NATIONAL ORGANIZATIONS

The American National Standards Institute (ANSI). ANSI, headquartered in Washington, DC with its operation center located in New York City, is a federation founded in 1918 that has been the administrator and coordinator of the U.S. private-sector voluntary standardization for 88 years. It is a private, not-for-profit organization supported by a number of organizations (i.e., its membership) in the public and private sectors. Today ANSI has approximately 1000 members from companies, organizations, government agencies and institutions.

ANSI's primary goal is "the enhancement of global competitiveness of U.S. business and the American quality of life by promoting and facilitating voluntary consensus standards and ensuring their integrity." ANSI also "promotes the use of U.S. standards internationally, advocates U.S. policy and technical positions in international and regional standards organizations, and encourages the adoption of international standards as national standards where these needs meet the needs of the user country." ANSI is the U.S. member of ISO and, via the United States National Committee (USNC), the IEC.

ANSI does not itself develop standards but functions, rather, as a coordinating body for the purpose of encouraging development and adoption of worthwhile standards as American National Standards. It looks to its organizational members, as well as to other concerned organizations, for accomplishing the task of standards development. These development activities may be performed wholly within one of these organizations, or in accredited committees organized and administered by one of these organizations and operating under a set of rules meeting the basic procedures of ANSI.

A large number of standards (more than 10,000 in 2002) that are processed for adoption by ANSI are developed, approved, and published by any of more than 270 standards development organizations accredited by ANSI. Organizations wishing to become ANSI accredited must develop and maintain procedures (subject to periodic audit by ANSI) that consistently adhere to a set of ANSI requirements or procedures entitled "ANSI Essential Requirements: Due Process Requirements for American National Standards*," that govern the consensus development process. Three of the more commonly used methods are described below:

1. *Accredited organization method*. The organization method is most often used by associations and societies that have, among other activities, an interest in developing standards. Although participation on the consensus body is open to all interested parties, members of the consensus body often participate as members in the association or society. The organization's procedures must

*Available for download from the ANSI Web site.

meet the general requirements of the ANSI procedures. By choosing to use this method, flexibility is provided, allowing the standards developer to utilize a system that accommodates its particular structure and practices.

2. *Accredited standards committee method*. Accredited standards committees are standing committees of directly and materially affected interests created for the purpose of developing a document and establishing consensus in support of this document for submittal to ANSI. The committee method is most often used when a standard affects a broad range of diverse interests or where multiple associations or societies with similar interests exist. The committee serves as a forum where many different interests, without a common membership in an organization or society, can be represented. Accredited standards committees are administered by a secretariat. An accredited standards committee develops and maintains its own operating procedures consistent with the requirements of the ANSI Procedures.
3. *Accredited canvass method*. A standards developer using the canvass method identifies, to the extent possible, those who are directly and materially affected by the activity in question and conducts a letter ballot or canvass of those interests to determine consensus on a document. Although canvass developers provide ANSI with internal procedures used in the development of the draft American National Standard, the due process used to determine consensus begins after the draft standard has been developed. Standards developers using the canvass method must use procedures consistent with the requirements of the ANSI procedures.

Figure 28-5 is an organization chart for ANSI. Standards work is under the Executive Standards Council (ExSC). The ExSC has responsibilities for American National Standards and U.S. participation in those international standards activities in which ANSI participates. ExSC is also responsible for maintaining ANSI procedures for standards development, development and coordination of U.S. position in international standards activities, and establishing and supervising such groups as needed to carry out these responsibilities. The approval and withdrawal of American National Standards has been delegated by the ExSC to the Board of Standards Review (BSR), one of the boards that report to the ExSC.

The Association for Electrical, Electronic and Information Technologies (VDE). VDE, located in Frankfurt am Main, Germany, is a nonprofit, nongovernmental organization founded in 1893 and is one of the largest technical and scientific organization in Europe with more than 33,000 members. In the area of standards in electrotechnology, VDE participates with The German Institute for Standardization (DIN) on the German Commission for Electrical, Electronic and Information Technologies of DIN and VDE (DKE), in both the IEC and CENELEC. The DKE is a joint organization of VDE and DIN; however, VDE is responsible for the administration of DKE. Approximately 4500 persons are involved in the standards process in more than 300 committees and subcommittees, and 300 working groups. The focus of standardization has shifted in the recent past toward the development of European Standards (based on international standards where possible). Pure national standards activity has dropped to approximately 5% of all work done.

VDE also performs certification and testing through the VDE Testing and Certification Institute, which was originally established in 1920 as the VDE Test Center. The VDE Testing and Certification Institute works together with other organizations, such as UL in the United States and CSA in Canada, in approximately 50 countries.

Standards Australia (SA). SA, located in Sydney, Australia, is an independent not-for-profit organization whose primary role is "to prepare Australian Standards through an open process of consultation and consensus in which all interested parties are invited to participate." SA was originally founded in 1922, as the Commonwealth Engineering Standards Association. In 1929, it became the Standards Association of Australia and was granted a Royal Charter in 1950. Its name officially changed to Standards Australia in 1988. SA represents Australia on both the ISO and the IEC. It is also involved in regional standardization activities and was a founding member of the Pacific Area Standards Congress (PASC). It maintains close ties with Standards New Zealand (SNZ) and has a formal agreement with SNZ for developing and publishing joint standards.

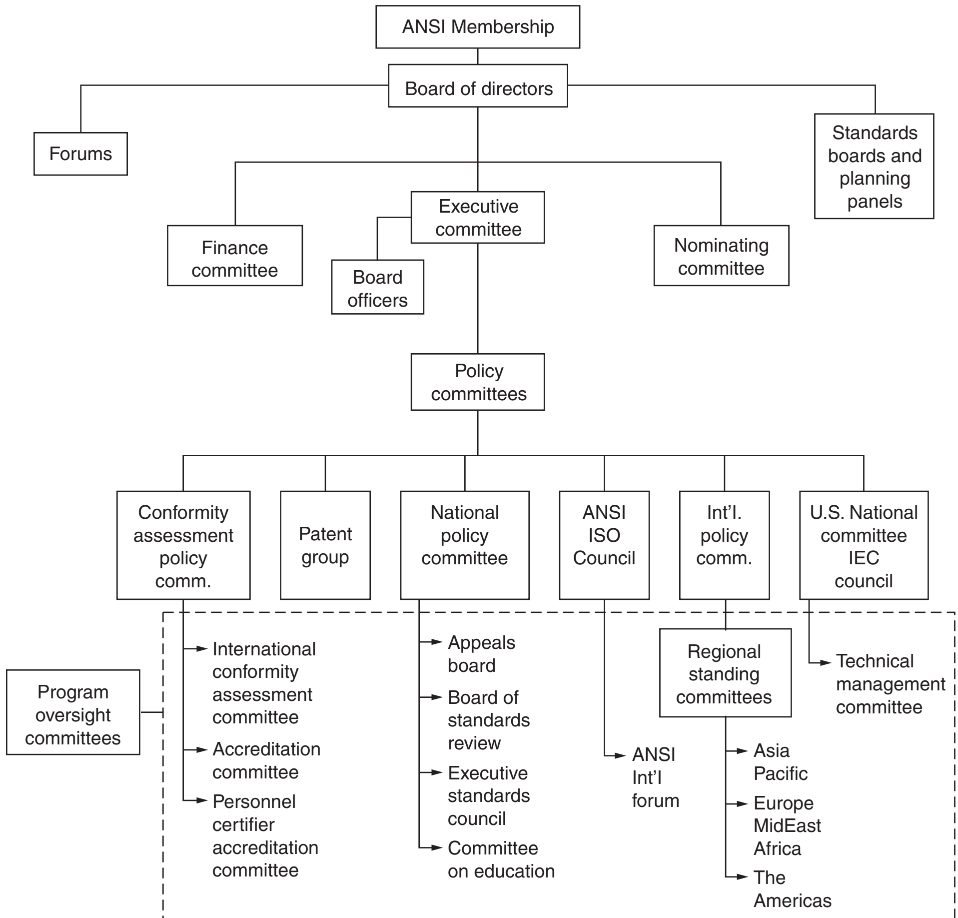


FIGURE 28-5 ANSI organization chart.

At present, approximately 9000 volunteer experts serve on 1500 committees that are responsible for the development and maintenance of approximately 6400 Australian Standards. SA has a subsidiary company, Quality Assurance Services Pty. Ltd., which provides system and product certification services.

Standards Council of Canada (SCC). The SCC, located in Ottawa, Canada, is a federal crown corporation with a mandate to "coordinate and oversee the efforts of the National Standards System, which includes organizations and individuals involved in voluntary standards development, promotion and implementation in Canada." The SCC reports to Parliament through the Minister of Industry. The Governing Council of the SCC consists of 15 members representing a broad spectrum of stakeholder interests. The council is responsible for setting the strategic direction of the organization, ensuring the fulfillment of the SCC's mandate and providing direction on governance matters. Canadian participation in the ISO and the IEC operates under a program under the SCC. The SCC acts as the secretariat for the Canadian National Committee of the ISO (CNC/ISO) and the Canadian National Committee of the IEC (CNC/IEC).

There are more than 15,000 Canadian volunteers involved in standardization activities and some 400 organizations have been accredited by the SCC, as SDOs, certification organizations, testing and calibration, laboratories, registration organizations or auditor certifiers and trainers. Of these organizations, four have been accredited as SDOs. These are Bureau de normalization du Québec (BNQ), Canadian General Standards Board (CGSB), Canadian Standards Association (CSA), and Underwriters' Laboratories of Canada (ULC). These accredited SDOs may submit standards they develop to the SCC for approval as National Standards of Canada (NSC). To become an NSC, a standard must be developed by a process incorporating the fundamental principles, be subject to public scrutiny, be consistent with or incorporate appropriate international and national standards, and be available in the English and French languages. A NSC cannot in any way be presented in such a manner that it will act as a barrier to trade.

28.10 OTHER STANDARDS DEVELOPERS

The Institute of Electrical and Electronics Engineers—Standards Association (IEEE-SA). The IEEE-SA, located in Piscataway, N.J. (U.S.A.), has been responsible for all matters related to standards since its formation in 1998. Actually, the concept of the IEEE-SA was approved by the IEEE Board of Directors in 1996 as part of an overall restructuring of IEEE as it moved into the twenty-first century. Prior to 1998, the IEEE standards program was administered by a Standards Board that reported to the Board of Directors. Under the IEEE-SA, the Standards Board still exists; however, it reports to the IEEE-SA Board of Governors. Standards development in the IEEE, which is the world's largest professional society, with more than 365 000 members in 150 countries, goes back to 1884, when the AIEE began to develop standards. Even today, the IEEE publishes 25% of all of the world's literature in electrotechnology.

The IEEE-SA was formed "to provide a major entity that would offer increased responsiveness to the standards interests of IEEE societies and their represented industries." The IEEE-SA provides services and innovative standards development processes that will keep pace with the needs of users. New classes of membership in the IEEE-SA have been approved, such as corporate membership, which will assist the IEEE-SA in realizing its strategic objectives. At present, the IEEE-SA has more than 30,000 volunteers and more than 70 corporate members participating in its standards activities. The IEEE-SA develops standards in such diverse subjects as broadcasting and communications, electrical practices for large industry (mining, textiles, shipbuilding, transportation, cement plants, and others), instrumentation and measurement, insulators and insulation, magnetics, motors and generators, nuclear power, power apparatus and systems, recording, symbols and units, electrical transmission and distribution, medical devices, electrical practices for commercial buildings and hospitals, emergency power systems, stationary batteries, and information technology. The IEEE-SA also serves as secretariat for a number of accredited standards committees (ASCs), an example of which is ASC C2, the *National Electrical Safety Code*® (NESC®).*

The National Electrical Manufacturers Association (NEMA). NEMA, located in Rosslyn, Va. (U.S.A.), is responsible for the development and maintenance of over 500 standards. NEMA is the largest trade organization for manufacturers of electrical products in the United States, and its 400 member companies are domestic firms varying in size from small companies to large diversified companies. It develops standards in the technical committees of its nine divisions covering

**National Electrical Safety Code*® is a registered IEEE trademark, and its acronym, NESC® is a registered IEEE service mark of the IEEE.

products in such fields as building equipment, power electronics, industrial electrical equipment, insulation, lighting, power equipment, wire and cable, radiation imaging products, and industrial automation.

NEMA technical committees comprise engineers designated to represent member companies who are manufacturers of electrical equipment. Since manufacturers are most knowledgeable in the technology associated with their respective products, NEMA committees are highly competent in developing product standards that realistically take into consideration the economic tradeoffs that are essential to practical standardization. Standards are adopted by consensus with final approval given by the NEMA Codes and Standards Committee.

NEMA Standards are generated in four classifications:

1. NEMA Standard—defines a commercially standardized product subject to repetitive manufacture.
2. Suggested Standard for Future Design—suggests an approach to future product improvement or development.
3. Authorized Engineering Information—included as part of other NEMA standards to explain data or information.
4. Official Standards Proposal—proposed draft for adoption by some other organization such as ANSI.

NEMA is a member of and actively participates in ANSI. It administers the work of a number of ASCs, acting as the secretariat.

The National Fire Protection Association (NFPA).† The NFPA, located in Quincy, Mass. (U.S.A.), has been active in standards development since its founding in 1896. Working under the direction of its Standards Council, the technical committees of NFPA, composed of organization representatives, personal members, and liaison members from other technical committees, develop standards documents that are then subject to public review as a result of advance publication, and are finally approved at a semiannual meeting of the entire NFPA membership. Although dedicated to fire prevention, NFPA is responsible for a series of electrical standards, the most noted of which are the *National Electrical Code* (NFPA 70) and the *Life Safety Code* (NFPA 101).

The NEC covers "the installation of electric conductors and equipment in public and private buildings or other structures (including mobile homes, recreational vehicles, and floating buildings), industrial substations, and other premises (such as yards, carnivals, and parking lots). The NEC also covers installation of optical fiber cable." The NEC is adopted and enforced in all the 50 U.S. states. It is also the basis for electrical codes in several other countries. For example, the Venezuelan electrical code is the 1981 edition of the NEC. The first edition of the NEC was published in 1897. The NFPA took over responsibility for the NEC in 1911. The NEC is produced and maintained by volunteers on 20 codemaking panels and an 11-member correlating committee that oversees the work of the panels.

The *Life Safety Code* "provides minimum requirements for the design, operation, and maintenance of buildings and other structures for safety to life from fire and similar emergencies." The *Life Safety Code* has been adopted by many states. Additionally, all health-care facilities in the United States receiving Medicare or Medicaid funding must comply with the *Life Safety Code*. The groundwork for the *Life Safety Code* began in 1913, when the NFPA appointed its first "Committee on Safety to Life." Its first pamphlet was on exit drills in factories, schools, department stores, and theaters. By 1921, the committee had produced a more comprehensive *Building Exit Code*. The *Life Safety Code* is produced and maintained by members of 15 technical committees.

†*National Electrical Code*®, *Life Safety Code*®, and NEC® are registered trademarks of NFPA.

Underwriters' Laboratories, Inc. (UL).[††] UL, located in Northbrook, Ill, (U.S.A.) is the leading third-party certification organization in the United States and the largest in North America. It is an independent, not-for-profit organization, founded in 1894, whose mark is recognized throughout the world as a symbol of safety.

UL maintains and operates laboratories for testing devices, systems, and materials with relation to public safety. Products so tested and meeting its requirements are eligible for UL "listing." UL maintains an inspection and follow-up program in factories where UL-listed devices are manufactured. UL representatives conduct in-factory and in-the-field inspections of manufacturers' procedures for assuring production compliance with UL requirements. Such requirements appear in appropriate UL Standards for Safety which are developed by UL under procedures that involve consultation with industry and government experts and consumers, among others.

The majority of insurance underwriters in the United States, and many federal, state, and municipal authorities, either accept or require listing or classification by UL as a condition of their recognition of devices, systems, and materials having a bearing upon life and fire hazards.

UL is divided into several engineering departments, each dealing with distinct and separate subjects as follows: electrical heating, air conditioning, and refrigeration; casualty and chemical hazards; burglary protection and signaling; fire protection; and marine. Each department has prepared standards for systems, materials, and appliances.

UL publishes (lists) the names of companies who have demonstrated the ability to provide products conforming to its requirements. Listing authorizes the manufacturer to use the laboratories' listing mark (classification marking, recognition marking, or certificate) on the listed products. UL submits its standards to ANSI for adoption as American National Standards.

UL is active in international standards development and product certification, and is a registrar for quality systems for various systems, including ISO 9000. UL acquired the Danish National Testing and Certification Organization (DEMKO) in July 1996 from the kingdom of Denmark, making it a wholly owned subsidiary of UL. This acquisition provides direct European certification of products to customers worldwide.

UL is responsible for the development and maintenance of more than 800 safety standards, the evaluation of more than 17,000 products, and more than 2500 quality system registrations. In 1998, UL estimated that approximately 14 billion products enter the marketplace each year bearing a UL mark.

28.11 U.S. GOVERNMENT REGULATORY STANDARDS BODIES

The U.S. Consumer Products Safety Commission (CPSC). The CPSC is an independent federal regulatory agency that was created by Congress in 1972, in the Consumer Products Safety Act. The mission of the CPSC is to "protect the public against unreasonable risks of injuries and deaths associated with consumer products." The CPSC has jurisdiction over approximately 15,000 consumer products. Exceptions are motor vehicles, trucks, and motorcycles, which are covered by the Department of Transportation; foods, drugs, and cosmetics, which are covered by the Food and Drug Administration; and alcohol, tobacco, and firearms, which are covered by the Department of the Treasury.

The CPSC works to reduce the risk of death and injury to consumers by developing voluntary standards with industry, and issuing and enforcing mandatory standards. If no feasible standard would adequately protect the public, the CPSC can ban a product. It can also require a recall of products to arrange for their repair. The CPSC also conducts research on potential product hazards and sponsors consumer educational activities (e.g., through the media). All work related to standards or regulations are published in the *Federal Register*. Initially, notice of the proposed

[††]Underwriters' Laboratories, Inc.® and its acronym, UL®, are registered trademarks of Underwriters' Laboratories.

standards or regulation is published for comment. Final regulations are also published in the *Federal Register*.

The U.S. Environmental Protection Agency (EPA). The EPA is an independent federal agency created by Congress in 1970. Its mission is "to protect human health and to safeguard the natural environment—air, water, and land—upon which life depends." The EPA develops regulations and has additional responsibilities for enforcement of these regulations.

The EPA, like other U.S. government agencies, follows a prescribed rulemaking process. The U.S. Congress establishes these requirements to support the development of quality rulemaking and protect the rights of those affected by the rules. Examples of these requirements include the Administrative Procedure Act, the Regulatory Flexibility Act, the Unfunded Mandates Reform Act, the Paperwork Reduction Act, and the National Technology Transfer and Advancement Act. The five stages of rulemaking are

1. *Prerulemaking*—actions to determine whether the agency should initiate rulemaking
2. *Proposed rules*
3. *Final rules*
4. *Long-term actions*—prerulemaking, proposed rule and finals rules expected to be published beyond the next 12 months
5. *Completed actions*—actions that are promulgated and published or actions that are no longer being considered

Comments from interested parties are encouraged during the rulemaking process. All rulemaking activities are published in the *Federal Register*.

The U.S. Federal Communications Commission (FCC). The FCC is an independent U.S. government agency established in 1934 by Congress in the Communications Act of 1934. Its mission is "to regulate interstate and international communications by radio, television, wire, satellite and cable." The FCC's jurisdiction is the 50 states, the District of Columbia, and U.S. possessions. The FCC is organized in bureaus and offices, and the official statement of the FCC actions is called an *order*. Most of the FCC's documents are issued by seven major regulatory bureaus or offices. When the FCC considers a change to its regulations, it issues a Notice of Proposed Rule Making. Essentially all documents issued by the FCC since 1994, can be found on-line through its website.

The U.S. Federal Trade Commission (FTC). The FTC enforces federal antitrust and consumer protection laws and seeks to ensure that the nation's markets function competitively, and are vigorous, efficient, and free of undue restrictions. The FTC has enforcement and administrative responsibilities under 37 separate acts.

The U.S. Food and Drug Administration (FDA). The FDA is one of the oldest federal agencies involved with the protection of consumers. In the area of electrotechnology standards, it is involved with setting standards for medical devices and radiation-emitting products. This work is done under the FDA's Center for Devices and Radiological Health (CDRH). The CDRH is responsible for "ensuring the safety and effectiveness of medical devices and eliminating unnecessary human exposure to man-made radiation from medical, occupational, and consumer products." Products covered in the CDRH's scope include thousands of medical devices (e.g., pacemakers), videodisplay terminals, microwave ovens, medical x-ray machines, and medical ultrasound devices.

National Institute of Standards and Technology (NIST). NIST, which is under the U.S. Department of Commerce, has been involved in the development of standards since its founding in

1901 (then known as the National Bureau of Standards). The documents it produced played a great role in the industrial development of the United States in a broad range of industries, including steel manufacturing, railroads, electric power, and telephone communications. Today it is involved in manufacturing partnerships, quality systems, measurement and testing laboratories, and advanced technology programs. In quality systems, it sponsors the Baldridge National Quality Program and the Malcolm Baldridge National Quality Award.

NIST also develops standards and guidelines for federal computer systems. These documents, known as *Federal Information Processing Standards* (FIPSs), are approved by the Secretary of Commerce. FIPSs are generally developed only when there are no acceptable industry standards available or there are compelling federal requirements (e.g., security). There are also Federal Standards (FED-STDS) developed for telecommunications; however, these are developed by the National Communications System (NCS) and are approved by the General Services Administration (GSA).

The major focus of NIST is on information technologies. Under the National Technology Transfer and Administration Act of 1995, NIST supports the development of national and international voluntary standards as the preferred source of standards for use by the federal government. NIST cooperates with national and international standards organizations, trade associations, consortia, user groups, and so forth to have needed standards developed.

The U.S. Nuclear Regulatory Commission (NRC). The NRC is an independent federal agency established by the Congress under the Energy Reorganization Act of 1974, to "ensure adequate protection of the public health and safety, the common defense and security, and the environment in the use of nuclear materials in the U.S." Its predecessor was the U.S. Atomic Energy Commission. NRC regulations are issued under the U.S. *Code of Federal Regulations* (CFR) Title 10, Chapter 1. Typically, rulemaking is initiated by the NRC's staff; however, any member of the public may petition the NRC to develop, modify, or rescind any regulation. During the rulemaking process, the documents are published in the *Federal Register* and interested parties are allowed at least one opportunity to comment. In some instances, the NRC will hold meetings and workshops before a proposed rule is drafted to obtain a broad range of input from interested parties. Most often in the area of electrotechnology, the NRC will endorse existing industry standards by issuing a regulatory guide (RG). The RG will describe methods that are acceptable to the NRC for applying the standard. The NRC may also use the RG to make recommendations or guides presented in a standard mandatory requirements. The NRC also issues NUREG reports to provide information and expertise to support the NRC's decision making and to assess potential technical issues.

The U.S. Occupational Safety & Health Administration (OSHA). OSHA, which is under the U.S. Department of Labor, was established by Congress in 1970 under the Occupational Safety and Health Act to "save lives and prevent injuries and illnesses in American workplaces." OSHA has responsibilities for developing standards and regulations, and enforcement. OSHA may set standards on its own initiative or in response to petitions from state and local governments, employers, labor representatives, standards organizations, or any other interested party. If it is decided to develop a standard, any one of several advisory committees may be tasked with the development. The process is very similar to that for other U.S. government agencies, including notice and comment period for interested parties.

28.12 CONTACTING STANDARDS ORGANIZATIONS

There are literally thousands of organizations around the world that are involved in standards development. Today, the World Wide Web (WWW) makes contacting these organizations easier than ever. The following listing is just a small sample of what is out on the Web. Many of the websites listed contain links to the websites of other standards developers.

International

IEC (International Electrotechnical Commission)
3, rue de Varembé
PO Box 131
CH-1211 Geneva 20
Switzerland
http://www.iec.ch

ISO (International Organization for Standardization)
ISO Central Secretariat
1, rue de Varembé
Case postale 56
CH-1211 Geneva 20
Switzerland
http://www.iso.org

JTC1 (Joint Technical Committee 1 Information Technology)*
http://www.jtc1.org

ITU (International Telecommunication Union)
Place des Nations
CH-1211 Geneva 20
Switzerland
http://www.itu.ch

ISOC (Internet Society)
1755 Wiehle Ave.
Suite 102
Reston Va. 20190-5108
U.S.A.
http://www.isoc.org
(At this Website the information on the following related Websites can be found.)

IETF (Internet Engineering Task Force)
http://www.ietf.org

IESG (Internet Engineering Steering Group)
http://www.ietf.org/iesg.html

IAB (Internet Architecture Board)
http://www.iab.org

IRTF (Internet Research Task Force)
http://www.irtf.org

IANA (Internet Assigned Numbers Authority)
http://www.iana.org

CIE (International Commission on Illumination)
Kegelgasse 27
A-1030 Wien (Vienna)
Austria
http://www.cie.co.at/cie

Regional Organizations

CEN (European Committee for Standardization)
Central Secretariat
36, rue de Stassart
B-1050 Brussels, Belgium
http://www.cenorm.be

ETSI (European Telecommunications Standards Institute)
650, route des
Lucioles
F-06921 Sophia-Antipolis Cedex France
http://www.etsi.org

CENELEC (European Committee for Electrotechnical Standardization)
35, rue de Stassart
B-1050 Brussels
Belgium
http://www.cenelec.org
[At this Website a considerable number of contacts are provided for other regional organizations such as the European Computer Manufacturers Association (ECMA), as well as the national standards bodies for member countries such as the Electro-Technical Council of Ireland (ETCI), and affiliate countries.]

*JTC1 is a joint committee of both the ISO and the IEC. The ITU has an official liaison to the committee.

Selected National Organizations

Australia
SA (Standards Australia)
http://www.standards.org.au

Canada
CSA (Canadian Standards Association)
5060 Spectrum Way, Mississauga, On L4W 5N6
http://www.csa.ca

SCC (Standards Council of Canada)
270 Albert St., Suite 200
Ottawa, ON. K1P 6N7
Canada

CNC/IEC (Canadian National Committee of the IEC)

CNC/ISO (Canadian National Committee of the ISO)
http://www.scc.ca

TSACC (Telecommunications Standards Advisory Council of Canada)
http://www.tsacc.ic.gc.ca

France
UTE (Union Technique de l'Electricité)
http://www2.ute-fr.com

AFNOR (L'Association FranÇaise de Normalisation)
http://www.afnor.fr

Germany
DIN (German Institute for Standardization)
http://www2.din.de

VDE (The Association for Electrical, Electronic and Information Technologies
DKE (German Commission for Electrical, Electronic and Information Technologies of DIN and VDE)
http://www.vde.com/vde_en

Italy
CEI (Comitato Elettrotecnico Italiano)
http://www.ceiuni.it

United Kingdom
BSI (British Standards Institution)
BEC (British Electrotechnical Committee)
http://www.bsi-global.com

United States

ANSI (American National Standards Institute)

USNC (U.S. National Committee of the IEC)

25 West 43rd Street, 4th Floor
New York, NY 10036
U.S.A.
http://www.ansi.org

(At this Website a considerable number of links are provided to Websites for standards-developing organizations located in the United States.)

ASC T1 (Accredited Standards Committee T1—Telecommunications)

T1 was retired in 2004 and its work assumed by ATIS.
http://www.atis.org

INCITS (International Committee for Information Technology Standards)

formerly Accredited Standards Committee X3 and NCITS.
http://www.incits.org

Selected Standards-Developing Organizations

AAMI (Association for the Advancement of Medical Instrumentation)
http://www.aami.org

ABMA (American Bearing Manufacturers Association)
http://www.abma-dc.org

AEIC (Association of Edison Illuminating Companies)
http://www.aeic.org

AHAM (Association of Home Appliance Manufacturers)
http://www.aham.org

AIIM (Association for Information and Image Management-The EMC Association)
(formerly National Microfilm Association)
http://www.aiim.org

ANS (American Nuclear Society)
http://www.ans.org

API (American Petroleum Institute)
http://api-ep.api.org

ARINC (Aeronautical Radio, Inc.)
http://www.arinc.com

ASA (Acoustical Society of America)
http://asa.aip.org

ASQ (American Society for Quality)
(formerly American Society for Quality Control)
http://www.asq.org

ASTM International (formerly American Society for Testing and Materials)
http://www.astm.org

ATIS (Alliance for Telecommunications Industry Solutions)
(formerly Exchange Carriers Standards Association)
http://www.atis.org

EEI (Edison Electric Institute)
http://www.eei.org

EIA (Electronic Industries Alliance)

ECA (Electronic Components, Assemblies, Equipment and Supplies Association)
www.ec-central.org

GEIA (Government Electronics & Information Technology Association)
www.geia.org

JEDEC (JEDEC Solid State Technology Association)
http://www.jedec.org

TIA (Telecommunications Industry Association)
http://www.tiaonline.org
http://www.eia.org

IS Alliance (Internet Security Alliance)
www.isalliance.org

NSTEP (National Science and Technology Education Partnership)
www.nationalstep.org

Electro (Electro-Federation of Canada)

CAMA (Canadian Appliance Manufacturers Association)

CEASA (Canadian Electronic and Appliance Service Association)

CEMC (Consumer Electronics Marketers of Canada)

CEMRA (Canadian Electrical Manufacturers Representative Association)

EEMAC (Electrical Equipment Manufacturers Association of Canada)

MIISC (Medical Imaging and Information Systems Council)

S&D (Supply and Distribution Council
http://www.electrofed.com

IEEE (Institute of Electrical and Electronics Engineers)
http://www.standards.ieee.org

IESNA (Illuminating Engineering Society of North America)
(formerly Illuminating Engineering Society)
http://www.iesna.org

ISA (The Instrument, Systems and Automation Society)
(formerly Instrument Society of America)
http://www.isa.org

IMAPS (International Microelectronics and Packaging Society) (formerly the International Society for Hybrid Microelectronics)
http://www.imaps.org

ITI (Information Technology Industry Council)
[formerly Computer and Business Equipment Manufacturers Association (CBEMA)]
http://www.itic.org

NACE (NACE International)
(formerly National Society of Corrosion Engineers)
http://www.nace.org

NEMA (National Electrical Manufacturers Association)
(includes standards of the Insulated Cable Engineers Association, ICEA)
http://www.nema.org

NETA (International Electrical Testing Association)
(formerly National Electrical Testing Association)
http://www.netaworld.org

NFPA (National Fire Protection Association)
http://www.nfpa.org

SAE (SAE International)
(formerly Society of Automotive Engineers)
http://www.sae.org

SEMI (Semiconductor Equipment and Materials International)
http://wps2a.semi.org

SMPTE (Society of Motion Picture and Television Engineers)
http://www.smpte.org

UL (Underwriters' Laboratories, Inc.)
http://www.ul.com

VITA (VMEbus International Trade Association)
http://www.vita.com

U.S. Government

CPSC (Consumer Products Safety Commission)
http://www.cpsc.gov

EPA (Environmental Protection Agency)
http://www.epa.gov

FCC (Federal Communications Commission)
http://www.fcc.gov

FDA (Food and Drug Administration)
http://www.fda.gov

FTC (Federal Trade Commission)
http://www.ftc.gov

NIST (National Institute of Science and Technology)
http://www.nist.gov

NRC (Nuclear Regulatory Commission)
http://www.nrc.gov

NTC (National Telecommunications System)
http://www.ncs.gov

NTIA (National Telecommunications & Information Administration)
http://www.ntia.doc.gov

OSHA (Occupational Safety and Health Administration)
http://www.osha.gov

Other Organizations

CITEL (Inter-American Telecommunication Commission)
Executive Secretariat of CITEL Organization of American States
1889 F Street NW
Washington, DC 20006
U.S.A.
http://www.citel.oas.org

WIPO (World Intellectual Property Organization)
34, chemin des Colombettes
PO Box 18
CH-1211 Geneva 20
Switzerland
http://www.wipo.int

WTO (World Trade Organization)
Rue de Lausanne 154
CH-1211 Geneva 21
Switzerland
http://www.wto.org

The World Wide Web Consortium (W3C)
In the Americas, contact W3C at:
Massachusetts Institute of Technology
32 Vasser Street, Room 32-G515
Cambridge, MA 02139
USA
http://www.w3.org

Webpages of Interest to Those Searching for Standards-Related Information

NSSN (National Standards System Network)
(a free online information service providing bibliographic information for more than 225,000 approved standards)
http://www.nssn.org

WSSN (World Standards Services Network)
(a network of publicly accessible servers of standards associations around the world)
http://www.wssn.net

New Approach Standardisation in the Internal Market
(a web site sponsored by CEN, CENELEC, ETSI, the European Commission and EFTA.)
http://www.newapproach.org

University of Waterloo—Canada: standards and specifications written by scholarly societies
(provides links to the Internet sites of many standards organizations)
http://www.lib.uwaterloo.ca/society/standards.html

U.S. General Services Administration—Index of U.S. Federal Specifications, Standards and Commercial Items
(an alpha numerical listing of federal specifications and standards, including a separate listing of all canceled or superseded federal specifications and standards)
http://apps.fss.gsa.gov/pub/fedspecs/index.cfm

U.S. Government Printing Office—Federal Register Database
(provides a database for rules, proposed rules, and notices of federal agencies and organizations, as well as Executive Orders and other presidential documents in ASCII or PDF format for 1995 to present (the 1994 federal register is also available, but it has no fields or section identifiers)
http://www.gpoaccess.gov/fr/index.html

USNRC Rulemaking Forum
(provides information on the NRC's rulemaking process, current rulemaking information and documents, draft regulatory guides, and NUREGs*
http://ruleforum.llnl.gov/index.html

World Wide Legal Information Association
(provides information on product standards and the law)
http://www.wwlia.org

*NUREGs are U.S. NUclear REGulatory reports normally followed by a number that makes the identification of the NUREG unique, e.g., NUREG 0700 is Guidance for Control Room Design Reviews.